THE UNIVERSITY OF ALBERTA

AN EXPERIMENTAL INVESTIGATION OF VOWEL ALTERNATION IN ENGLISH

by

ⓒ RESTITUTO MENDOZA CENA

A THESIS
SUBMITTED TO THE FACULTY OF GRADUATE STUDIES AND RESEARCH
IN PARTIAL FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN
PSYCHOLINGUISTICS

DEPARTMENT OF LINGUISTICS

EDMONTON, ALBERTA
FALL, 1976

THE UNIVERSITY OF ALBERTA

FACULTY OF GRADUATE STUDIES AND RESEARCH

The undersigned certify that they have read, and recommend to the Faculty of Graduate Studies and Research, for acceptance, a thesis entitled "An Experimental Investigation of Vowel Alternation in English", submitted by Restituto Mendoza Cena in partial fulfilment of the requirements for the degree of Doctor of Philosophy in Psycholinguistics.

Gary D. Prideaux
Supervisor

Lois Marckwardt

John Hogan

....

....
External Examiner

Date September 20, 1976

dedication:

to estelita alano cena,
who has been assiduously preparing
my breakfast cereal.

ABSTRACT

Vowel alternation in English was subjected to an experimental test for psychological reality. It was hypothesized that if the pattern of vowel alternations in English facilitates the learning of novel forms exemplifying correct English alternations and if it interferes in the learning of novel forms exemplifying alternations not found in English, then it is psychologically real. Five correct English alternations (e.g., ay/I) and 20 non-English alternations (e.g., ay/e), which served as control conditions, were embedded in made-up adjective noun pairs, e.g., subk[ay]pe/subk[I]pity, malp[ay]de/malp[e]dity, etc. The word pairs were presented aurally to 40 university students as paired-associates to be learned, using the study-test technique.

The results showed that, in all three learnability indices used, the class of correct English alternations was learned significantly more readily than the class of non-English alternations, and, furthermore, that knowledge of these correct English alternations interfered in the learning of the non-English alternations. It was concluded that the vowel alternation generalization in English is psychologically real.

In an attempt to reconcile the positive results of the experiment with the negative results of two previous productivity studies on English vowel alternation, a distinction is proposed between psychological reality and productivity, whereby a productive process is considered to be necessarily psychologically real, but that a psychologically real process may not be strong enough to be productive. A scale of "psychological productivity" is proposed; in this scale, surface phonetic processes would rate high, and purely morphological, non-phonetically conditioned processes would gravitate toward the lower end.

ACKNOWLEDGEMENTS

This thesis is an attempt to exemplify a way of thinking about language. For this frame of mind, and for the many suggestions that led to the development and improvement of the thesis, I wish to thank my supervisor, Dr. Gary D. Prideaux, and the other members of the examining committee: Dr. John T. Hogan, Dr. Lois Marckworth, Dr. Bernie Rochet, and Prof. William S-Y Wang. I am particularly indebted to Dr. Prideaux for his patience in straightening my convoluted thoughts. He has been a constant source of encouragement.

I also wish to thank Prof. Wm. J. Baker, Dr. Bruce L. Derwing, and Dr. Anton Rozsypal for many helpful suggestions.

My fellow graduate students in the department helped at every stage in the preparation of the thesis. I am especially grateful to John Gray, with whom I have had many hours of profitable discussion of practically every aspect of this thesis, from designing the experiments to text formatting. Pat Ford recorded the stimulus materials for the pilot study, and Tracy Love Derwing the stimulus materials for the actual study. Richard Stephens taught me little tricks on how to tame the PDP 12. Geoffrey Rowe helped in the analysis of data. Terry Hull made the final copy much more readable than it could possibly be in my hands. I am also indebted to Allan Opperthauser for his assistance in the technical aspects of stimuli preparation and presentation. I thank you all.

I must express my gratitude to three friends in the Philippines, who not only introduced me to the fascinating field of linguistics, but who also gave me numerous chances to advance my knowledge in this area: Profs. B. P. Sibayan, F. T. Otanes, and E. P. Dagot.

My years of study in linguistics have been supported by the Ford Foundation, the American Council of Learned Societies, and the University of Alberta.

Needless to say, any errors of thought and style in the thesis remain the responsibility of Zachariah.

TABLE OF CONTENTS

CHAPTER ONE: BACKGROUND TO THE STUDY

An Overview of the Study 1

The Theoretical Motivation

The Need for External Validation of Rules 3
Linguistic Data 13

The Descriptive Framework

Productive and Nonproductive Rules 18
Vowel Alternation in English 22

Some Methodological Considerations

Formal and Substantive Reality 27
Data-Collection in Experimental Phonology 29

CHAPTER TWO: REVIEW OF RELATED STUDIES

The Role of Linguistic Experience 36

Testing for the Reality of Representations

On the Reality of MSC's 49
On the Psychological Manifestations of MSC's 52
An Experiment on an Abstract Segment 56
The Reality of Intermediate Representations 61

Experiments on Phonological Rules

On the Generativity of PR's 64
Experiments on English Stress 68
Experiments on Vowel Alternations in English 76
Summary 82
The Experiment 84

CHAPTER THREE: METHOD

Subjects 88

Acquisition Test

Materials 88

Procedure 91

Design 92

Transfer Test

Materials 94

Design and Procedure 94

Post-Test Interview 95

CHAPTER FOUR: RESULTS

Tabulation of Raw Scores 97

A Preliminary Overall Analysis 99

Analysis of the Learning Scores 103

Analysis of the English Interference Scores 106

Another Source of Interference 109

Analysis of the Transfer Scores 114

CHAPTER FIVE: DISCUSSION

On Two Sources of Interference

Morphological Interference 117

Phonetic Interference 117

Relevance to Two Models of Verbal Learning

Stimulus-Response Models 124

Information-Processing Models 128

On the Organization of the Lexicon 133

On the Issue of Psychological Reality 145
A Note on 'Simplicity' and Learnability 152
Addendum: Other Explanations of the Results 155

CHAPTER SIX: SUMMARY AND A SUGGESTED FURTHER STUDY

Summary 158
A Suggested Further Study 161
REFERENCES 167
APPENDIX A: Stimulus Materials 181
APPENDIX B: Subject Instructions 182
APPENDIX C: Tables 187
APPENDIX D: Figures 199

LIST OF TABLES

Table 1. Frequency Data, Summed over Subjects and Trials 187

Table 2. Summary Table, Stimulus x Actual Response 188

Table 3. Summary Table, Required Response x Actual Response 188

Table 4. Overall ANOVA Results 189

Table 5. Results of Post-Hoc Comparisons among Selected Categories 190

Table 6. ANOVA Results, Correct Test Scores 192

Table 7. ANOVA Results, First Correct Trial Scores 192

Table 8. ANOVA Results, Last Error Trial Scores 193

Table 9. Correct English Scores as Percent of Total Error Scores 193

Table 10. ANOVA Results: Correct English/Correct Test 194

Table 11. Error Scores Not Attributable to Correct English, Summarized as a Function of Stimulus and Actual Response 195

Table 12. Error Scores not Attributable to Correct English, Summarized as a Function of Required Response and Actual Response ...195

Table 13. Results of Discriminant Function Analysis of Correct Test Scores 196

Table 14. ANOVA Results, Transfer Scores 198

APPENDIX D
FIGURES

Figure 1A. Interaction between Stimulus and Actual Response, Front Formation Correct Test Scores 199

Figure 1B. Interaction between Stimulus and Actual Response, Back Formation Correct Test Scores 200

Figure 2A. Interaction between Stimulus and Actual Response, Front Formation First Correct Trial Scores 201

Figure 2B. Interaction between Stimulus and Actual Response, Back Foration First Correct Trial Scores 202

Figure 3A. Interaction between Stimulus and Actual Response, Front Formation Last Error Trial 203

Figure 3B. Interaction between Stimulus and Actual Response, Back Formation Last Error Trial 204

Figure 4A. Cell Deviation Scores of the Categories Correct Test and Correct English, Front Formation 205

Figure 4B. Cell Deviation Scores of the Categories Correct Test and Correct English, Back Formation 206

Figure 5. Front Formation Error Scores not Attributable to Correct English 207

Figure 6A. Interaction between Stimulus and Actual Response, Transfer Test, Front Formation 208

Figure 6B. Interaction between Stimulus and Actual Response, Transfer Test, Back Formation 209

CHAPTER ONE

BACKGROUND TO THE STUDY

> A simple analysis might be better than a complex analysis, but only if it is true that speakers would account for the data by means of the simpler analysis. By just looking at static data, there is no way at present for a linguist to determine what regularities speakers will capture.-- R. Skousen (1972, p. 567)

An Overview of the Study

The present study is an experimental investigation of the psychological reality of vowel alternation in English. By "psychological reality of vowel alternation" is meant the relevance to certain mental activities of the generalization underlying vowel alternations.

In recent years, the notion of psychological reality has played a prominent role in discussions of rule validity. Linguists who have seriously considered the question of the validity of linguistic rules invariably inquire about the psychological reality of the rules in question. Even though in many cases these inquiries serve merely to illustrate a point and are never really pursued experimentally, the recognition of a psychological dimension in rule validation is a move in the right direction. Since language underlies a certain mental reality, linguistic concepts ought to be characterizable in mental terms. This is a basic requirement that must be imposed on a mentalistic linguistic theory such as generative linguistics.

Accordingly, the approach to the problem of rule evaluation pursued here is experimental. A two-part learning experiment has been carried out to determine the relevance of a phonological generalization to the mental activity of learning. The view is taken that only if a linguistic generalization does have behavioral correlates can the generalization be said to underlie a certain mental reality, and that only those rules based on psychologically real generalizations can be said to properly belong to a mentalistic grammar.

Interest in this line of research is twofold: First, it is important to know that a linguistic generalization proposed for a particular language is in fact a generalization that speakers actually do (or do not) employ. Second, experimental data of the sort reported here provide a basis for the more general problem of the external validation of formal linguistic concepts.

The Theoretical Motivation

On the Need for External Validation of Rules

The classical generative view of theory validation may be summarized as follows: Given a collection of objects, there is an n-number of ways of describing the system of relationships among the objects. Subsets of this number of ways are adequate descriptions depending on certain increasingly stringent formal requirements of adequacy. From among the many observationally adequate grammars, only a subset are descriptively adequate, and from this set of descriptively adequate grammars, the most stringent set of adequacy requirements selects only one explanatorily adequate grammar. A rule that fits into this explanatorily adequate description is considered to be valid. Among proponents of this theory, an unduly simplistic view has prevailed that all that needs to be done to advance linguistic research is to constrain the formal requirements of the theory of grammars such that it can decide from among the descriptively adequate grammars that which is explanatorily adequate.

A major shortcoming of this view is its dependence on the assumption that the transformational methodology can in practice discover many descriptively adequate grammars. In actual fact, not one descriptively adequate grammar of any language has so far been formulated. There are, therefore, no descriptively adequate grammars to choose from. This failure to produce even one descriptively adequate grammar may be symptomatic of a basic error. And the error may in fact be that linguists have looked long enough at linguistic objects as abstract mathematical objects, separate from the purpose for which they have been created, and separate from the human mind which created and use them.

This is not to deny the role of formalization in discovering interesting aspects of language. By winnowing out the grain from the chaff, that is, by abstracting significant generalizations from particular cases, a formal system gives an unobstructed view of the organization of the system. It thus provides, in linguistics, a language-independent framework for describing and comparing other linguistic systems, with the view towards discovering universal properties of linguistic systems. Furthermore, once the formal language is fairly developed, it can then feed on itself, and, in the manner of an axiomatic system, generate theorems about itself even prior to the consideration of empirical evidence. Formalization, thus, gives the linguist an opportunity to look at certain aspects of the inner mechanism of language.

Initial successes with the formal methodology have, however, biased the linguist's point of view against other linguistic methodology. While formal rules should serve only as one of the ways by which true generalizations about language may be discovered, formal rules now seem to have become the only path to linguistic enlightenment, and formal rules have thus become the ultimate objective of linguistic inquiry. Within this mold, linguistic inquiry stops after the formal expression of a generalization has satisfied the researcher, and suggestions to externally validate this formal statement are seldom entertained seriously. This enthusiasm for formalism may have been fanned by a hint from Chomsky that the mathematical underpinning of language is of a type hitherto unknown. About this, he wrote:

> ...the latter [grammatical transformations] being structure-dependent operations of a peculiar sort that have never been studied outside of linguistics, in particular, not in any branch of mathematics of which I am familiar. (Chomsky, 1969, p. 77)

If what is involved is a type of mathematical terrain yet uncharted, then the formal linguist can lay claim to certain liberties in the use of any formal devices and conventions he may find necessary in his search for an adequate description. Given this climate of opinion, the suggestion that when the formal description and external evidence clash, the formal description be upheld, was inevitable. Fodor and Garrett (1966) wrote that when an experimenter has uncovered negative evidence, he

> ...has not shown that the grammar is disconfirmed. A grammar is simply an axiomatic representation of an infinite set of structural descriptions, and the internal evidence in favour of the structural

> descriptions modern grammars generate is so strong that it is difficult to imagine their succumbing to any purely experimental disconfirmation. (p. 152)

Amidst a multitude of symbols and devices, and in the rush to find the neatest formulation, the purpose of linguistic inquiry and the nature of linguistic data appear to have been lost. It is true that searching for the simplest rule is an engrossing activity, but if such an activity is not related to basic questions of human nature, it reduces itself to the status of an intellectual exercise. Chomsky has stated what he feels the goal of linguistic inquiry should be. In discussing what a linguist does, he wrote:

> At the level of particular grammars, he is attempting to characterize knowledge of a language, a certain cognitive system that has been developed -- unconsciously, of course -- by the normal speaker-hearer. At the level of universal grammar, he is trying to establish certain general properties of human intelligence. Linguistics, so characterized, is simply the subfield of psychology that deals with the aspects of mind. (1968, p. 24)

Of necessity, then, the data of linguistic inquiry are psychological. Such data "consist[s] of the behavior of the speaker, listener, and learner of language" (Chomsky, 1959, p. 170).

If this view of the nature of linguistic inquiry is kept in mind, the insufficiency of formal criteria alone in evaluating linguistic solutions becomes quite obvious.

Herein lies a difference between a linguistic system and a truly abstract system. A description of a linguistic system must be founded on a certain external reality, the psychological reality of speech communication. The development of a truly abstract system, however, need not depend on the availability of evidence of external validity. Evaluative criteria for any linguistic statements, therefore, must relate to the nature of the human mind, what it can do, or what it actually does. This point has of course been expressed adequately by many linguists. For example,

> Chomsky (1967, p. 100):
>
> Sooner or later -- in some areas sooner, in other areas later -- it is going to be necessary to discover conditions on theory construction, coming presumably from experimental psychology or from neurology, which will resolve the alternatives that can be arrived at by the kind of speculative theory construction linguists can do on the basis of the data available to the linguists. That is, there will come a point, no doubt,...where one can set up alternative systems to explain quite a wide range of phenomena. One can think that this or that system is more elegant and much more deep than some other, but is it right?
>
> Derwing (in press):
>
> The fundamental problem which confronts scientific linguistic research today is not the formal one of deciding what kinds of rules can be formulated by the linguist within some particular arbitrary frame of reference, but rather the empirical one of discovering what kinds of rules are actually learned by the child in the acquisition of the language and ultimately actually employed by the adult in the production and comprehension of speech.

Ohala (1974b, p. 35):

> It is clear to an increasing number of linguists that the sterile, faddish formalism and unchecked speculation that characterizes most of generative phonology and its offshoot will not, by themselves, reveal anything about mental processes. Speculation must be guided and restrained by empirical evidence. Formalizations, if they are to be at all relevant, must have as their primary aim the formulation of theoretical statements in a manner that renders them more testable. Formalism and speculation if pursued as ends in themselves are only intellectual amusements.

The inadequacy of a purely formal criterion in evaluating competing proposals that are descriptively adequate has been demonstrated quite singularly in the failure of the simplicity criterion in resolving certain aspects of the abstractness controversy. The abstractness issue was raised by Kiparsky in 1968 (the paper was published in 1973; see Kiparsky, 1973a), in reaction to Chomsky and Halle's (1968) practice of postulating underlying phonological elements that never appear on the surface in any variants of the morphemes analyzed as having these underlying segments. The issue is whether to allow such representations. Abstract segments invariably require phonetically unmotivated rules, such as an absolute neutralization rule, which is postulated precisely for the purpose of preventing such abstract segments from appearing in surface phonetics. This point will be illustrated using Hyman's (1970) Nupe examples.

Nupe has a series of consonants that labialize before the back vowels [u, o], and palatalize before the front vowels [i, e]. Before the vowel [a], however, a three-way distinction among consonants occurs -- labialized [Cʷa], palatalized [Cʸa], and plain [Ca]. The descriptive problem is how to characterize these labialized and palatalized consonants. It is very natural for back (rounded) vowels and for front (unrounded) vowels to trigger labialization and palatalization, respectively. But for [a] to cause labialization as well as palatalization is irregular, and even more so since in other cases [a] causes neither labialization nor palatalization. Thus, the very natural labializing and palatalizing rules fly in the face of the labializing and palatalizing [a] which produces "exceptions". After considering a number of alternatives, Hyman opted for an abstract solution, one in which the labializing [a] derives from an abstract underlying back vowel /ɔ/, and the palatalizing [a] derives from an abstract underlying front vowel /ɛ/. Neither of these postulated segments appears in surface phonetics; where they occur in underlying structure, [a] appears in phonetic surface structure. Hence, an absolute neutralization rule must apply to neutralize the distinctions among these low vowels and to merge these vowels into the phonetic [a]. Thus, underlying /Cɔ/ and /Cɛ/ become [Cʷɔ] and [Cʸɛ] through labialization and palatalization, respectively, and the neutralization rule applies on these intermediate strings to produce the required [$C^{\omega}a$] and [$C^{y}a$].

It was the general thinking initially that solutions making use of abstract elements were more general, in effect, simpler. With two additional underlying segments and one rule, Hyman's solution not only accounts for a multitude of facts, but it does so in true generative fashion, and at the same time it preserves the exceptionality of two very natural, phonetically motivated rules. In contrast, a concrete solution to the Nupe problem is the postulation of a whole set of labialized and palatalized consonants that contrast with plain ones, an "uneconomical" solution if economy is to be measured in terms of size of the inventory of underlying segments (as Hyman pointed out, there are far more serious problems with this approach). There are better less abstract solutions (see, for example, Kiparsky, 1973a, Harms, 1973), but the point is that whichever solution is adopted, concrete or abstract, something is lost and something is gained. How effective is the formal measure of simplicity in evaluating this trade-off between loss and gain? In the case of the Nupe problem, how "simple" are two abstract segments and an absolute neutralization rule, as opposed to a host of non-abstract segments that "clutter-up" the inventory of underlying segments?

Crothers' (1971) reply was that, "it is not very meaningful to appeal to some prior notion of simplicity when we do not know to what extent imaginary [abstract] representations are in themselves non-simple" (p. CR-7). Kiparsky's objection to simplicity arguments is even more basic: "Simplicity arguments cannot justify absolute neutralization, for it is precisely the validity of simplicity arguments which is challenged" (1973a, p. 66). Hyman himself takes this same view. In his reply to Harms' (1973) criticism of his approach, Hyman (1973) states: "But economy is not the issue. As pointed out by Zimmer 1970, notions of simplicity have not helped in solving any critical issues in generative phonology" (p. 449).

Having rejected the relevance of the internal criterion of simplicity in resolving the issue, participants in the controversy, pro and con alike, turned to considerations external to the formal system. They appealed to substantive considerations either in terms of learnability or psychological plausibility. Thus, Kiparsky (1973a) expressed the view that a rule of absolute neutralization "is either impossible or hard to learn", and suggested that such a rule "be either excluded or specified as not highly valued" (p. 65). In Kiparsky (1974b), he expressed the view that the evaluation measure is not merely a simplicity metric, but, following Koefoed (1975), as "an explication of the concept 'learnability of the abstract system of a language'" (p. 328). Hyman, who defended the abstract position, reminded us that "our aim [in writing grammars] is not merely to find new and neater ways to present data, but to discover the nature of the mental mechanism activated by the child in language acquisition" (p. 59). His own justification for an abstract solution to the Nupe problem depended on sound changes in loan phonology. Crothers, who upheld the concrete position, wrote that "what is needed...is not more discussions of which solution is the simplest or the best, but more facts about how the speaker...actually handles the situation" (p. CR-8). To this Hyman must be said to have agreed, when, summarizing his reply to Harms (1973), he wrote that "considerations of economy and naturalness are poor substitutes for the psycholinguistic evidence which we sorely need" (p. 451).

It is evident that agreement has been reached to look beyond what Skousen calls "static data" (see quotation at the beginning of this chapter). Skousen's suggestion is a call to expand the data base to include the "dynamic data" of sound changes observable in the historical and ontogenetic contexts. Additional external data are available from other sources: sound changes observed in disturbed speech, in forced and manipulated speech in the experimental laboratory, in linguistic games in the form of street word-play or in the more elevated level of versification. Speech samples from these sources illustrate language in use--on the way to or departing from the standard language. Comparing the norm with departures from the norm can provide revealing insights on just those properties of the norm that are significant.

The issue of external validation of linguistic rules is related to the issue of the role of psycholinguistic experimentation in linguistic theory construction. The view taken in this study is that experimentation provides a way of substantiating linguistic speculations. Discussions of this issue abound in the literature, but, to my knowledge, it has not been approached in relation to the nature of linguistic data. The brief discussion that follows focuses on just this aspect of the issue.

Linguistic Data

A persistent criticism of experimental linguistics is that while competence is at issue, linguistic experiments operate in the realm of performance, and that experimental data are irrelevant in resolving issues in the area of competence. This criticism is based on a misconception of the nature of the data which all linguists, whether experimentalists or formalists, work with.

What is the nature of the data on which a formalist bases his conclusions? The establishment of the competence/performance distinction eliminates from grammar-writing errors attributable to performance factors like memory limitation, inattention, momentary failure to apply a rule, deliberate disobeyance of rules, etc. Certainly no one would wish to write a grammar based on utterances that native speakers, under better conditions, would consider ungrammatical. The distinction, therefore, ensures that the grammarian works on a sample of edited speech. But the data on which experimentalists base their inferences are also edited, in this sense, and are no more contaminated by performance errors than the set of data grammarians work on. In fact, the experimentalist's data may be much more strictly edited since he sees to it that what are reasonably extraneous factors are eliminated, or are properly controlled for. No experimentalist would let his data be distorted by, say, memory limitations, by requiring his subjects to repeat back a left branching sentence with twenty embeddings (unless this is what is at issue).

Furthermore, the criticism that experimentation is irrelevant to resolving issues in the area of competence relies on the unwarranted assumption that the subject in an experiment is incapable of achieving a state in which his performance matches his competence, but that the speaker writing a grammar is able to. This assumption is made because while the relevance to issues of competence of data supplied by a group of experimental subjects is doubted, the relevance of data generated by grammarians is never questioned in most of the literature on formal analysis (although particular judgments of individual linguists may, from time to time, be considered suspect). Both the grammarian writing a grammar of his language and the subject in an experiment are performing essentially the same task:

making judgments about their language. Why should one set of data be more valid than the other? Grammarians have not been trained to perform in an error-free manner, nor should their intuitions be any less subject to the constraints of sound data-collection procedure than anyone else's.

All experimentation requires data-collection of one type or another, and the linguist writing a grammar is essentially doing an experiment -- he formulates hypotheses and tests these hypotheses against data from his own intuition. It is in this sense that more credibility may be ascribed to an experimental grammar than to a formal grammar, since data collection in experimental grammar construction is governed by sound sampling procedures, whereas data collection in formal grammar writing is typically subject to a bias that has no control.

Thus, the critic of experimental linguistics is essentially making the claim that only grammarians are in a position to achieve a state of error-free performance, and, hence, only their judgments are relevant to specifying the nature of competence. Such a claim is rejected here. The intuitions of a non-grammarian are as relevant as the intuitions of a grammarian. They are, in fact, more relevant, since the grammarian's judgments are subject to systematic contamination in ways which those of the linguistically untrained subjects are not (for further discussion, cf. Derwing, 1973, chap. 7, sec. 3). Indeed, research in transformational grammar has a history of proposals and swift counter proposals, of vague visions and rapid revisions, of doubts, speculations, and total turnabouts, all of which point to the fact that at some point certain grammarians were in error and certain grammars were erroneous.

Furthermore, to this date no explicit account of how performance parameters actually distort statements about competence has appeared. In the absence of such an account, the criticism can be seen only as an unsubstantiated claim. In fact, some of the performance factors mentioned in the literature are more facetious than serious. For example, mention has been made of possible distortions of data due to food in the mouth, and even untimely death. To this problem, a colleague has offered an appropriate solution: we must, at all costs, avoid doing our experiment at mealtime, and never use as subjects patients in a hospital terminal ward.

An implication of the view that performance data are irrelevant to deciding issues of competence is that it insulates competence (=grammar) from experimental verification (Labov, 1971). This is an unfortunate consequence, considering that competence is meant to be interpreted as a reflection of the structure and organization of the human mind. If tests of performance cannot reveal competence, then, as Zimmer (1969) has pointed out, "the appeal to competence is really quite empty, since there is no independent support for ascribing to speakers some competence that is not revealed by their performance" (p. 320).

On the relation between competence and performance, the view adopted here is that of Steinberg (1970). Steinberg questioned Chomsky's assumption that competence is describable independent of performance considerations. He suggested that "Chomsky's rules of competence are indistinguishable, in principle, from rules of performance" (p. 180). To Steinberg, a competence grammar may differ from a performance grammar only in the way the rules are organized; he suggested that "rules of competence...are ordered for maximum utility with respect to the processes of production and understanding" (p. 190). The suggestion, in short, is that the formulation of models of competence be based on the processes of language use. Thus, models of competence should be evaluated in terms of the processes of production and comprehension.

In the next section, the descriptive framework of the present experiment is discussed. Following current practice, the terms "psychologically real" and "productive" will be used synonymously. In a later chapter, a distinction between these two important terms will be proposed. Descriptive aspects of productive rules will be surveyed, and on the basis of this survey, the question of the productivity of the English vowel alternation rules will be addressed

The Descriptive Framework

Productive and Nonproductive Rules

Productive rules were distinguished from nonproductive rules in terms of the opposition major/minor rules. Thus, Krohn (1972) suggested that major rules are productive while minor rules are less productive. Discussing some recent psycholinguistic validation experiments, Kiparsky (1973a) criticized these studies for not drawing a distinction between major and minor rules; he appeared to be making the suggestion that some of the rules that were shown to be experimentally nonproductive were nonproductive for just the reason that they were minor rules. There are, however, no generally agreed upon properties of minor and major rules. To Kiparsky, major rules are those which in the unmarked case apply and minor rules are those which in the unmarked case do not apply. To Krohn, a minor rule is one that makes use of rule features; his own formulation of the English vowel shift requires a plus rule feature, hence, in his view, vowel shift is a minor rule, and, therefore, a less productive rule.

Saltarelli (1972) drew a distinction between predictive and nonpredictive rules. To Saltarelli, predictive rules are formulations that define "an open (trans-finite) domain," whereas nonpredictive formulations define "a closed (finite) domain circumscribed to the existing lexicon" (p. 320). Predictive rules are "linguistically significant", as these rules "provide a principled explanation for the speaker's linguistic behavior" (p. 320), whereas nonpredictive rules "have no obvious significance as they characterize a finite list, which is definitely weaker, if not trivial, use of the notion of generalization" (p. 320). Although there is no direct equation of predictive=productive and nonpredictive=nonproductive, it is evident in the discussion that such an equation was intended. Thus, he wrote that "a speaker's knowledge of his language [represented by predictive generalizations] is best investigated in its productive derivational processes" (p. 339). Predictive rules, however, were equated with psychologically real rules. He wrote that nonpredictive formulations "are of no relevance to how the mind works, and can't be incorporated but provisionally in a synchronic grammar (p. 321). In another paper (Saltarelli, to appear), he viewed predictive rules to be similar to phonetic constraints, and proposed that "synchronically, only PC [phonetic constraints] can claim some kind of mental reality" (quoted by M. Wang, 1974, p. 12).

M. Wang (1972) established a distinction between last and nonlast rules. Last rules apply last, that is, they do not precede any other rule (type) in the grammar although they may be preceded by other (layers of) last rules. She proposed to view last rules as surface structure constraints, and nonlast rules as rules of the morphological component, which describe regularities of closed sets of lexical items. She claimed that "only those sound changes which correspond to last rules in the synchronic grammar can be productive, and that just these _must_ be productive" (p. 278). A further claim was made that the level of last rules is a significant level in generative phonology, and she offered as evidence the observation that hypercorrection and rule simplification occur at this level.

Skousen (1972) suggested that a rule is psychologically real or productive if it eliminates all possible exceptions that might arise, for example, by borrowing or by the operation of other rules. Thus, to Skousen, a rule is productive if children's speech errors tend to move in the direction of the rule, or if the rule operates in foreign borrowings or new word coinages. The converse claim, that a rule is not real if surface violations of the rule do not tend to become eliminated, has been criticized, however, as too strong by Kiparsky (1973b).

Kiparsky's (1973a) distinction between _opaque_ and _transparent_ rules should also be mentioned. Transparent rules are without exceptions, or have far fewer exceptions than opaque rules. Kiparsky suggests that transparent rules are easier to learn, hence, they are more highly valued. In essence, then, transparent rules are more productive than opaque rules. The opposition "learnable" vs. "nonlearnable" and "plausible" vs. "implausible" (Derwing, 1973) may be other distinguishing aspects of productive and nonproductive rules.

By far the most extensive discussion of formal productivity and phonological rules is in M. Wang (1974). The bulk of the discussion is on the formal ways linguists make unproductive processes appear productive or exceptionless. These devices include rule ordering, inclusion of morphological information in the structural description of rules, use of diacritic features, diacritic use of phonological features, and rule features. All of these devices insure that the rules involved apply just to those words that exhibit the sound changes being described, and, hence, the rules are made to look exceptionless. Wang concluded:

> ...we may use the presence of such devices [ordering and the use of nonphonological information in structural descriptions] as defining characteristics of nonproductive rules. Thus, a phonological rule may be recognized as nonproductive if it contains diacritic features or any reference to morphological information, including morpheme boundaries, or if it is in counter-feeding order to some other rule. Conversly, we may recognize a rule as productive if it has none of these characteristics. (pp. 69-70)

In summary, productive phonological rules exhibit the properties of major rules, last rules, and surface phonetic constraints; they are phonetically motivated rules, whose specifications require no nonphonological information; they apply without exceptions, and in open domains, and, hence, they are predictive; they are phonological rules as opposed to morphological rules. Nonproductive rules have the converse properties. We will see shortly that the rules that are the subject of the present study -- the English vowel alternation rules -- possess none of the properties that allegedly distinguish productive rules.

Vowel Alternation in English

The major motivation for the vowel alternation rules is provided by the alternations exemplified below:

Front Vowel Alternations:

ay/i	divine/divinity
iy/e	serene/serenity
ey/æ	profane/profanity

Back Vowel Alternations:

aw/ʌ	profound/profundity
uw/ɔ	school/scholar
ow/ɔ	verbose/verbosity

These alternations have been formalized in Chomsky and Halle (1968), Wang (1968), Ladefoged (1971), Krohn (1972), and Stockwell (1972). Since the present study is concerned with the question of the psychological reality of these alternations and not with the question of which formulation is the correct one, only a brief outline of one of the formulations, that of Chomsky and Halle, will be presented.

In the Chomsky-Halle formulation, the alternations were given the abstract underlying sources /i, e, æ, u, o, ɔ/, respectively (where the macron signifies tenseness). To derive the surface alternations from the postulated underlying segments, three main rules and two adjustment rules are required; the main rules are vowel shift, diphthongization, and trisyllabic laxing; these rules are here referred to collectively as the vowel alternation rules (hereafter, VAR).

In general, stressed tense vowels that agree in backness and rounding undergo shift in height. In the first stage, high vowels become mid and mid vowels become high. In the second stage, mid vowels become low and low vowels become high (for one-stage formulations of the shift rule, cf. Ladefoged, 1971, and W. Wang, 1968; Wang's formulation is discussed briefly in Chapter 6 of this thesis). However, in the presence of certain suffixes, for example, the noun-forming suffix "-ity", trisyllabic laxing applies. Trisyllabic laxing is ordered before shift, to prevent the crucial vowel in words like "divinity" from shifting.

If the environment for shift is restricted to tense vowels, it will not give the correct output for forms like profundity and scholar. Suffixation with "-ity" and "-ar" laxes the crucial vowels, thus producing the incorrect prof[u]ndity and sch[o]lar. To account for these forms, Chomsky and Halle extended the environment for shift to include lax $\underline{u}$ and $\underline{o}$. Thus, prof[u]ndity becomes prof[o]ndity, and a rounding adjustment rule adjusts $\underline{o}$ to $\underline{\wedge}$. This rounding adjustment rule, and a backness adjustment rule, also take care of certain dialectal differences in the surface forms, e.g., div[ay]ne-div[æy]ne, verb[a]sity-verb[ɔ]sity. All of these changes are summarized as follows (underscoring represents tenseness):

	Underlying Segments:						$\underline{i}$	$\underline{e}$	$\underline{æ}$	$\underline{u}$	$\underline{o}$	$\underline{ɔ}$
Diph/Lax	$\underline{i}$y	$\underline{e}$y	$\underline{æ}$y	$\underline{u}$w	$\underline{o}$w	$\underline{ɔ}$w	i	e	æ	u	o	ɔ
VS, Step 1	$\underline{e}$y	$\underline{i}$y		$\underline{o}$w	$\underline{u}$w					o		
VS, Step 2	$\underline{æ}$y		$\underline{e}$y	$\underline{ɔ}$w		$\underline{o}$w					ɔ	
Backness Adj	$\underline{a}$y											
Rounding Adj				$\underline{a}$w						^	a	a
Output	$\underline{a}$y	$\underline{i}$y	$\underline{e}$y	$\underline{a}$w	$\underline{u}$w	$\underline{o}$w	i	e	æ	^	a	a

On the basis of the formal characterization of the properties of productive rules outlined in an earlier section, we may now ask the question of whether vowel alternation in English is productive. Productive rules are exceptionless. Certainly VAR is not without exceptions: obesity (in most dialects), nicety, rarity, nudity, crudity, etc. In the sense of Kiparsky (1973a), VAR may be considered as opaque. VAR cannot be treated as an output or surface phonetic constraint since there are many instances of surface VGCV in English (h*i*bernate, *i*solate, pr*o*bity, v*a*gary, etc.), which should never occur if VAR functions as an output constraint. VAR applies in a closed lexical set, and, therefore, is not predictive outside of this set. VAR requires nonphonological information in its specification, i.e., the diacritic [+F]. This diacritic is attached to the vowel when it is in the nonphonological context PAST, a context required to cover alternations like *sit*-*sat*, *sing*-*sang*. "The great vowel shift," wrote M. Wang (1974, p. 43), "is an example of a nonproductive phonological process in English." Cearley (1974) considers VAR as a morphological rather than a phonological process, and makes the claim that vowel nonshift is regular and vowel shift is irregular:

> Morphological rules define classes of irregular forms....all forms subject to the application of a particular morphological rule are *marked* for its application....For instance, *profane* is lexically marked [+laxing]; when conjoined to the affix +*ity*, the stem is subject to the application of the laxing rule: prof[ey]ne-prof[æ]nity. Forms such as *obese*-*obesity*, *nice*-*nicety*, which do not show laxing, are not exceptions to the rule, but are regular: they do not show alternations! Those forms which are marked for the application of a morphological rule are by their very nature irregular. Given the tendency toward paradigmatic regularity, we should find that forms such as those just discussed tend to drain out of irregular paradigms, becoming completely regular. (p. 30)

Maher (1971) proposed to characterize the relationship between members of pairs like *divine*-*divinity* as etymological rather than synchronic, and, hence, the alternation need not be made part of the synchronic grammar. A similar view has been expressed by Oloffsson (1974). By all formal "standards", then, VAR is not a productive process, and, therefore, it should not be psychologically real.

No direct claim has been made that VAR is psychologically real. Its treatment as such is implicit in SPE, where it is considered as a very general phonological process, and in fact the "pivotal process of modern English phonology" (p. 102). An indirect claim of reality for this rule has been made by C. Chomsky (1970). Discussing vowel shift and the abstract representation of words undergoing shift, she writes:

> Are these abstract representations that are postulated by the linguist merely convenient fictions that the linguist manufactures for the purposes of his grammar, or do they have a psychological reality for the language user? In other words, is the claim that the orthography [=underlying lexical representation] corresponds to something real in the linguistic knowledge of the reader based on anything that the reader can honestly be said to know?
>
> It seems to me that in a very real sense the lexical level of representation and the corresponding aspects of English orthography do have a psychological reality for the language user. (p. 295).

It seems that if a claim is made for the psychological reality of the underlying representations that are input to a rule, then the rule for which these representations have been postulated must in some sense be considered to be likewise psychologically real. The question of the validity of VAR has thus been reduced to the question of its psychological reality or productivity.

Some Methodological Considerations

Formal and Substantive Reality

Since the present study is an inquiry into an aspect of the notion "psychologically real rule", it must be made clear what aspect of the rule under consideration is being tested here. A rule has two components: form and content. Therefore, in discussions of the psychological reality of rules, a distinction must be made between the reality of the content of a rule and the reality of its form (Kiparsky, 1968). Two rules that differ in form may have the same content, and thus account for the same phenomena. If the predictions turn out to be correct, then the rules have content reality. Such a demonstration, however, does not say which form or which rule has reality. Hence, testing for the reality of a rule involves testing for its form as well as its content; these tasks are, to a certain extent, independent of each other (this point was made earlier by J. Ohala, 1970, p. 5). Testing for content validity means testing for knowledge of the generalization underlying the rule, in whatever form the generalization may be represented, while testing for formal validity amounts to testing for the way this knowledge is represented in terms of the formal language of a particular theory. Thus, when Steinberg and Krohn (1975) wrote about the "psychological validity of Chomsky and Halle's vowel shift rule", it is presumed that what was meant is the content and not the form of the rule since the experiments they described tested only for speaker knowledge of vowel alternations in English.

Formal properties are of two types; one is specific to a rule and the other relates to formal properties shared by other rules. An example of the former is the property of the Chomsky-Halle vowel shift rule of being a two-step formulation; Wang (1968) has suggested a one-step formulation. Formal properties of the second type relate to the use of such abbreviatory devices as curly brackets, angled brackets, the neighborhood notation, and the like. Kiparsky (1968) is an attempt to adduce historial evidence in support of the 'psychological reality' of these devices. Moskowitz (1973) tested experimentally for the more rule-specific type of formal property.

Since many competing rules differ crucially only in their form, it is desirable that formal validity be tested experimentally. More basic than addressing the question of formal reality, however, is the task of determining first whether a rule has content reality, since it is pointless to examine the correctness of the form of a rule if the rule in the first place has no substantive basis. The present study is directed to this more basic task. The question asked is whether the generalization underlying the vowel alternation rule in English guides linguistic performance.

In discussing related studies in the next chapter, reference will be made to methodological aspects of these studies. It will be useful to mention at this point certain data-collection problems in experimental phonology.

Data-Collection Problems in Experimental Phonology

Problems of data-collection may arise as a result of errors committed in two aspects of the experiment, namely, the selection of materials and the selection of the task itself. In many cases, phonological experiments require that the stimulus materials be novel, yet similar to familiar ones in crucial and relevant respects. If the materials are too novel to the subjects, the subjects may not see the applicability on these materials of the generalization being tested. In which case, the results are irrelevant. On the other hand, if the subjects are too familiar with the materials, they may perform independently of the experimental variables being manipulated, in which case the results are uninterpretable. A balance must be maintained between novelty and familiarity.

For example, if one is testing for the predictions of the Chomsky-Halle stress rules for English, it is obviously unwise to present to the subjects real English words. In such a case, if the subjects did put stress in the right place, it could not be concluded with confidence that the subjects obeyed certain stress rules, since the correct stress could just as well be supplied without recourse to rules. That is, the stress pattern of a word may be, like its meaning, part of a set of information arbitrarily paired with the word, independent of any intrinsic property of the word or of any directing rule. This interpretation must be eliminated if the putative rule is to be tested. On the other hand, if a completely novel nonsense word is presented as stimulus, nonfulfilment of the predictions of the rule may be due not to a lack of knowledge of the rule, but to the fact that the subjects did not see the relevance of the rule being tested to the test word. There is evidence that foreign items are not readily integrated into the grammar of the borrowing language and, hence, are not always subject to its regular phonology (Kiparsky, 1973a).

A reasonable solution is the judicious use of nonsense words. The use of nonsense forms in testing for the predictions of a rule is a perfectly valid research strategy, as Ross (1972, p. 245) has noted. Nonsense words may be constructed such that they do not violate any structural constraints of the language of the subjects; in other words, they can be made to sound like native words. Experimentally, such forms are introduced as "possible" words of the language, and the subjects are asked to treat them as real words which they have not had the chance to meet before the experiment. Since the subjects have in fact not met these words before, any phonological properties these forms may exhibit in the experimental situation are not idiosyncratic properties of these words, but are general properties of the phonological system.

Berko's (1958) use of nonsense words in her study on the productivity of certain English morphological rules was motivated by this desire to balance novelty and familiarity of the materials. However, while she may have succeeded in making her stimulus materials sound Englishlike, the very Englishness of these materials may have engendered a related problem, namely, the problem of analogical patterning. The nonsense items may have provoked direct association with real words, and, hence, the subjects' responses to these nonsense items may have been responses to real word analogues. Berko's subjects may have said "two wug[z]", thinking perhaps of "two bug[z]". Anderson (1972) referred to unpublished work by Raymond Baird which suggested "that children at one stage can form plurals to nonsense-syllables, properly choosing from [-z, -s, -Iz], only if they knew a real word rhyming with the nonsense item" (p. 420). In other words, the subject may not be using a rule to direct him to produce a form; he may be merely patterning his responses after known forms. In Kiparsky's words, "the fact that a speaker treats made-up words in accordance with an existing putative rule does not immediately justify the reality of the rule: for example, an analogical pattern of some kind may be operating, which would have rather different properties from the supposed rule" (Kiparsky, 1973a, pp. 101-2).

The term "analogy" as used in this study will be made clear. Kiparsky did not explain what he meant by "analogical pattern"; and neither did Chomsky when he objected to the use of the notion of analogy as a basis for the explanation of the comprehension of novel sentences (Chomsky, 1968, p. 30). One, however, gets the impression that they intended performance by analogy to mean a kind of performance that is not mediated by generative rules. This is essentially Hankamer's (1972) interpretation of Chomsky's notion of "derivatively generated" (as opposed to "directly generated by the grammar"), which is a process of formation by analogical rule, a type of rule that is neither phrase structural nor transformational. It is in this nongenerative sense that the term is used in this study, and its use as such does not involve a commitment by the present writer; the intention is to show that the effect of this type of analogy in experimental phonology can be appropriately controlled for.

Analogy, of course, has a respectable status in the structuralist tradition. Paul's (1891) and Bloomfield's (1933) use of the term has been interpreted by Derwing (1973, p. 309) to be synonymous with, or an instance of, linguistic creativity. Perhaps Kiparsky was referring to the "proportional interpretation" of analogy (e.g., sister:sisters = brother:brothers), which, in another paper (Kiparsky, 1974), he pointed out to be inadequate in explaining analogical changes. This use of "analogical phonological rule", Ohala (1974b) says "requires explicit reference to existing words which manifest a certain pattern" (p. 36). But even in this sense, analogy is not totally un-creative; the fact that a subject can make a correct analysis of the sound properties of the new form, and on the basis of precisely the relevant characteristics, he is able to compare it with a member of the right class of forms, suggests a kind of linguistic ability that cannot in the least be considered to be superficial. Ohala (1974a,b) and Hsieh (1975) have in fact suggested that the role analogy plays in language learning may be more significant than hitherto realized by transformationalists.

The possibility of analogical patterning becomes all the more evident if we consider the type of experimental task commonly employed. The subject is presented with forms combinable according to certain English word-derivational rules. For example, he is given a base and a noun-forming suffix. He is asked to form a noun from these materials, and he is asked to pronounce the combination, typically in a sentence frame. The subject is given unlimited time. Whatever the subject's response, the experimenter acknowledges it. In this procedure, analogical patterning may creep in on two fronts. First, the subject may analogize with real word exemplars. That is, he may consult his internal lexicon and match the nonsense word with a real word. Second, he may use his early responses as models for subsequent responses. If his early responses are acknowledged by the experimenter as acceptable, the subject may advert to whatever strategy he may have used in processing the early materials and use it consistently thereafter, regardless of the nature of the succeeding materials.

Knowledge of linguistic rules is by and large 'hidden' even to the speaker himself, and one of the ways this hidden knowledge may be exposed is to make the task sufficiently demanding such that coping with the task is possible only if recourse to this hidden knowledge is made. Thus, the information contained in an incoming signal may be deliberately set to a level that is not up to threshold so that the subjects, in attempting to supply the 'missing parts', will draw on this hidden knowledge. Or the input information may come in chunks too large for immediate processing so that the subjects, in transforming the input signal to manageable units, will use language-based processes. The point is that, since psychological experimentation involves the quantification of behavioral manifestations of the mechanisms of the mind, the selection of the right measuring device (the experimental task) plays a crucial role in the outcome of the experiment. If the measuring device is not adequately calibrated, irrelevant results may be obtained.

Other procedural errors may be specific to particular experiments. Kiparsky (1973a) mentions some of the more obvious errors which have been made in recent experiments on English phonology:

> Rules which are restricted to the Romance part of the vocabulary (such as t --> s) are tested with Germanic morphemes; no distinction is drawn between major rules and minor rules, that is, between rules which in the unmarked case respectively do and do not apply; morphologically deviant words are made up as tests for rules (for example, subjects are asked to add deadjectival suffixes to nouns (e.g. maze + ity), and Romance suffixes to Germanic vocabulary). (p. 102)

In the second half of the next chapter, some of the more recent phonological construct validation experiments will be reviewed in the light of the points raised in this section.

CHAPTER TWO

REVIEW OF RELATED STUDIES

The Role of Previous Linguistic Experience

A basic assumption in linguistic psychology is that linguistic behavior is rule-governed. Underlying the use of language are rules that guide performance. Rejection of this assumption renders uninteresting the field of linguistic inquiry, for then it must abandon its search for underlying organizing principles. But more important, such a rejection implies acceptance of the counter-assumption that language acquisition and use are essentially unmediated by principles that make the job easier to accomplish. It is implausible that the human mind will opt for an essentially unorganized system when some organization is possible.

The assumption that linguistic behavior is rule-governed gives rise to the corollary assumption that if in fact speakers have internalized a process or rule, it will be used in situations that are new but in essence similar to those under which the generalization has been extracted. This is the basis for the experimentalist's expectation that the effects of certain of the unobservable mechanisms and entities postulated by particular theories may somehow be exposed. If these postulated mechanisms and entities have observable or measurable effects, then they may be psychologically credible. Stated in the context of a verbal learning experiment, "if the linguistic units have psychological reality and if the linguistic units are present in the material to be learned, then learning scores should vary in some way with the linguistic units" (Cofer, 1969, p. 342, in reference to the basic assumption of Johnson, 1964). One might say that this assumption underlies all linguistic construct validation experiments. What is the basis for this assumption?

The basis is the massive body of evidence on linguistic interference in the learning of verbal materials in experimental situations. In a review of these studies, Cofer (1969) concluded:

> It is virtually impossible to design units which are independent of S's language skills and background. The prominence we have given to meaningfulness, familiarity, association value, pronunciability, inter- word associations, sequential organization, category membership, and grammatical structures makes this abundantly clear. Past experience with the language results in skills and habits which interact, in an inexorable way, with contemporary verbal learning tasks. (p. 364)

The nature of this interaction, Glanzer (1967) suggested, is in the form of proactive interference, where previous linguistic experience inhibits or facilitates the learning of the experimental tasks.

That the subjects' linguistic experience interacts with the verbal learning task has been recognized since the beginning of verbal-related experimentation. Thus, Ebbinghaus' avoidance of real word materials in his memorization experiments was an attempt to control for word meanings and word interrelation. This led him to use nonsense syllables, on the assumption that nonsense syllables were equally meaningless and that if they associated with one another at all, any association is as strong as any other association, an assumption that Ebbinghaus himself realized as untenable.

In the last two decades, considerable progress has been made in identifying the variables that cause interference. There are two major research orientations in this area. The first deals with properties of verbal materials as discrete linguistic elements: meaningfulness, pronounceability, familiarity, frequency, concreteness , vividness and imagery content, orthographic distinctiveness, etc. Studies along this line have periodically been reviewed, and the more recent ones are Mandler (1967) and Cofer (1969, 1971).

The other line of research, and the one that has received more attention in the last decade, concerns properties of linguistic elements in sequence, i.e., structural and organizational properties. This shift in interest may be attributed to the rise to prominence of two conceptual models -- the cognitive approach to learning and generative linguistics. The cognitivist provides a view of the learner as a dynamic participant in the learning process, always in search of structures, strategies, and rules. The generativist provides a specification of linguistic structures whose interrelations are expressed in the form of processes or rules. These linguistic structures and processes are specified with a degree of precision that, in many cases, render them available for experimentation. The bulk of this latter type of research has been on syntax-semantics, and adequate reviews of these studies are available (cf. Glanzer, 1967; Bever, 1968; Fillenbaum, 1971; Gough, 1971; Johnson-Laird, 1974; Fodor, Bever, and Garrett, 1974).

The general picture that has emerged from these syntactic-semantic studies is that linguistic behavior, like any other type of cognitive activity, opts for the efficiency provided by a rule-governed system. Since syntactic-semantic generalizations and phonological generalizations are of the same theoretical status, our expectation is that phonological generalization be likewise a variable in mental activities.

In the remainder of this chapter, a selective review of studies in experimental phonology will be presented. Two considerations have guided the selection of studies in this review. First, the object of a study must be phonological in nature (as opposed to phonetic). This excludes the many studies on the multidimensional scaling of phonetic features, and studies on the role in learnability and recall of phonetic features (e.g., Wickelgren, 1965, 1966; Jenkins, Foss, & Greenberg, 1968; Klatt, 1968; Coltheart & Geffen, 1970). Second, the study should be experimental; this excludes studies on the psychological reality of certain phonological constructs using as bases for discussion normative/clinical data from disturbed speech (Schnitzer, 1972, 1974), non-pathological speech errors (Fromkin, 1973), allegro speech (Zwicky, 1972), linguistic games (Sherzer, 1970), writing systems (Bugarski, 1970; Scholes, 1972), and metrics (Kiparsky, 1972).

One of the earliest psycholinguistic experiments was conducted by Esper in 1925. He compared the learnability of three artificial mini-languages that varied in their degree of approximation to English. One of the languages that had to be learned departed from certain syntactic and phonological structural properties of English (a confounding that unfortunately was not controlled for). The phonological distortion consisted of an unnatural syllable break. Thus, the first CV of a CV#CCVC compound word had a referent distinct from that of the remaining CCVC segments. The CC clusters used in the experiment do not occur in English in syllable-initial position (e.g., lg, gd, zg, mb, etc.), and although the whole compound word was pronounceable, the part corresponding to the second referrent was not pronounceable due to the unpronounceability of the syllable cluster. As expected, this language was found harder to learn than the one that did not violate any structures of English.

The relevance of this experiment to the present study stems from the possibility that the phonological anomaly of the unnatural syllable break--or the unpronounceability of the initial syllable clusters--had interfered in learning. While the increased difficulty of this language may be said to be the confounded effect of both syntactic and phonological distortions, there is reason to believe that the phonological distortion contributed a great deal of the errors subjects made. For example, substitution errors showed that the subjects extended the first word of the compounds to include the next consonant, resulting in a more natural CVC.CVC division. In the words of Esper, "in general, we find a tendency [to] modify the non-English syllable divisions of Experiment II in accordance with English speech habits" (p. 38). It would seem that the phonological distortion of the sort mentioned above made the correct responses difficult to integrate, thereby rendering them less available as response items.

The phonological fusion experiments, initiated by Day (1968), and pursued extensively at Haskins Laboratories, have yielded interesting data on phonological mediation in the perception of the temporal order of phonemes in clusters. The technique requires presentation of dichotic pairs, e.g., *pay/lay*, one word to each ear, and the subject is asked to report what he hears. Day (1968) reported that subjects typically heard *play*. There was fusion when both stimulus words were real words, e.g., *pay/lay* --> *play*; when the stimulus items were nonwords and the fusions were real words, e.g., *banket/lanket* --> *blanket* (Day, 1970); and even when both stimulus and responses were nonwords, e.g., *gorigin/lorigin* --> *glorigin* (Day, 1968). Fusion was observed even when presentation of *banket* was delayed by up to 100 msecs (Day, 1970). In other words, even when the crucial sounds [b] and [l] were objectively ordered as [lb], subjects heard them as [bl]. This result was duplicated when the dependent measure was judgment of the temporal order of the crucial sounds. In the visual mode, comparable results were reported by Rommetveit (1968, p. 99), in which *shar/shap* were reported to have been seen as *sharp*. This subjective re-ordering of sounds in perception is reminiscent of the general finding of the click experiments (Ladefoged and Broadbent, 1960; Fodor and Bever, 1965), which showed that subjects subjectively relocated objective clicks towards major constituent boundaries (for a comprehensive review of the linguistic and nonlinguistic variables in fusion, see Cutting, 1975).

In explaining phonological fusion, Day (1970) suggested that, in the perception of linguistic materials, a "linguistic filter" is operating. Thus, when the subject 'perceives' lbanket, the filter rejects It and alters it to the acceptable blanket. Day writes:

> Before a subject can give a response, the results must be related to past experience with the language, perhaps by way of a linguistic filter or similar device. The filter operates on the basis of the sequential dependencies of phonemes in the language. (p. 84)

The particular constraint that is operating here is the restriction on the order of stop-liquid sequence in word initial position in English, which disallows *[lb] but not [bl]. That a linguistic explanation is in order is given some support in another experiment (Day and Cutting, 1970), in which the stimuli consisted of nonlinguistic materials. In this experiment, subjects made accurate judgments of the temporal order of signals. Furthermore, when the stimuli were reversible clusters, as for example, [sp] and [ps] in final position, re-ordering of sounds was not observed. That is, given the dichotic pair lis-lip, the subjects reported hearing lisp when lis led, and they reported hearing lips when lip led. Phoneme sequence misperceptions did not occur since no phonotactic constraint restricts the ordering of the sounds [s] and [p] in word final cluster.

The "verbal transformation" experiments (see Goldstein & Lackner, 1973) provide another line of support for the template role of natural language habits in the perception and learning of linguistic materials. In the earliest of these experiments, Warren and Gregory (1958) reported a tendency by subjects to hear different words from an unchanging stimulus presented repeatedly over and over; the phenomenon has been referred to as verbal transformation effect. The finding from this line of experimentation that is relevant to the present study is the observation that subjects' reports show evidence that phonological resegmentations and reorganization of the stimuli are subject to the phonological constraints of the language of the subjects. Thus, the sequence "tress-tress-tress" was reported to have been heard as "stress-stress-stress" (Warren and Warren, 1970). Even when subjects reported hearing nonsense words, these words tended to be segment sequences permissible in English (Warren, 1961). A similar conclusion was reached by Moorehead (1970), who wrote that "whenever a semantic or phonetic change was made, [his subjects] changed the original sequence to English in the most economical way according to the rules of English phonology" (p. 62).

The technique has great potential in psycho-phonological experimentation. Two experiments that employ this technique in directly testing for certain issues in generative phonology are Goldstein and Lackner (1973), who looked at the possibility that phonetic coding may be in terms of matrices of distinctive features (see, also, Lackner & Goldstein, 1975); and Moorehead (1970), who employed the technique to see if the phonological rules or strategies children use in processing non-English phoneme sequences are different from those used by adult speakers. Use of this technique may reveal something about consonant assimilation or permissible initial consonant clusters in English. Thus, if "tress-tress-tress" is transformed into "stress-stress-stress", would, say, "dress-dress-dress" be misperceived as "*sd*ress-*sd*ress-*sd*ress", and would "fress-fress-fress" be misperceived as "*sf*ress-*sf*ress-*sf*ress"?

Certain linguistic habits that influence learning may be more deep-seated than habits acquired through speaking a specific language. These habits may be conditioned by universal properties of languages. An experiment by Schane, Tranel, and Lane (1974) shows evidence that certain linguistic universals may be psychologically real.

Observing the formal distinction between natural and unnatural rules, Schane, Tranel, and Lane asked the question: Would an artificial mini-language system exemplifying a natural rule induce more learning than one that exemplifies an unnatural rule? An example of a natural rule is one that deletes a consonant before a word beginning with a consonant; the converse rule, one that deletes a consonant before a vowel, is unnatural (the basis for this distinction is the observation that the preferred syllable sequence structure in many languages is CVCVCV...). A rule of this sort has many exemplars in French. Thus, *petit* 'little' + *garsõ* 'boy' is *peti garsõ*, (cf. *petit ami* 'little friend'). Schane, Tranel and Lane claimed that it has very limited application in English; the a/an alternation is one example that is true of all dialects of English. They suggested that such a limited application does not necessarily mean that the structure CVCVCV is not preferred by English speakers, but that it may simply be the case that it is an accident of linguistic development that English, unlike French, does not draw heavily upon this linguistic universal. Thus, if it is in fact the case that this formal universal of preferred syllable structure has psychological reality, the prediction is that, given the chance, speakers of English will show preference for such a structure.

Schane, Tranel, and Lane first taught their subjects four novel nouns as paired associates to their English "translations". Two of the novel nouns began with a consonant sound, and the other two began with a vowel sound. In the second stage of the experiment, three novel adjectives, all with final consonant sounds, were paired with each of the novel nouns, following the natural rule (consonant deletes before consonant) for one group of subjects, and the unnatural rule (consonant deletes before a vowel) for the other group of subjects. The twelve phrases were cued by their English translations, and the subjects responded with the phonologically correct sequence of adjective and noun. The results showed that "the Ss learning the unnatural corpus had a strong tendency to give natural responses, whereas the converse was not true. Consequently they made many more errors en route to mastery than their natural counterparts,.... It appears that our Ss had implicit knowledge of the natural rule, even though it does not operate to any significant extent in English" (p. 351).

Why should phonological constraints of the sort involved in the experiments reviewed so far interfere in mental activities? The answer is not hard to see if these constraints are viewed as "expectations," "strategies," or "rules" that subjects have or use in handling linguistic materials. After all, the subjects go into the experimental laboratory with a lifetime of linguistic experience behind them. When they are asked to process materials that appear to them to be organized the way a human language is organized, they invariably employ modes of processing that they have become accustomed to use in their own language.

Psychologists have for some time now been talking about this type of proactive interference. Interference may be negative, in which case processing of the new materials is hindered, as in the Esper and the fusion experiments. Or it may be positive, in which case learning of the new materials is facilitated (cf. studies on syntactic facilitation, for example, Glanzer, 1962; Baker and Prideaux, 1975). Interference is explained in terms of such familiar psychological notions as response competition, response mediation, response similarity, and, lately, response restriction (Bower, 1970). When two response systems are sufficiently similar, learning or recall of the correct response is hampered by the fact that something similar to it competes for attention. Of these competing responses, the stronger response is invariably selected; the stronger response is that which is more strongly linked to the stimulus.

For example, in the context of the transfer schema A-B, A-C, if B is a type of response associated with linguistic usage, and C is the experimentally prescribed response, B interferes in the learning of C and this interference is a function of the similarity of B to C. In the testing phase, response B may predominate over C if C has not been fully learned. Imperfect learning of C may arise if interference by B is sufficiently strong, in other words, if the linkage between A and B has not been effectively extinguished. Thus (all other things equal), the dominance of the linguistically determined response B is taken as an indirect evidence for the existence of a linguistic mediator between A and B. Bower (1970) restated this general idea by saying, in effect, that the predominance of B over C may result if a mediating rule exists that restricts the range of possible responses to A to responses of the type B. Thus, a subject's recognition that a rhyming rule, for example, relates stimulus and response helps him to restrict his choices of responses to words that rhyme with the stimulus word. In the present study, Bower's response restriction hypothesis is the basis of the expectation that the Vowel

Alternation Rules should facilitate the learnability of word pairs that exemplify the rules. This point is further discussed in the section on The Experiment at the end of this chapter.

Testing for the Reality of Representations

On the Reality of MSC's: Zimmer (1969)

Zimmer (1969) set out to ascertain the degree to which native speakers of Turkish were aware of the regularities expressed by certain Turkish morpheme structure conditions (MSC's). Awareness was to be construed as evidence for their reality. How does one go about testing for 'awareness' of this sort of knowledge? Zimmer reasoned that if subjects were presented with pairs of nonsense forms, one word of each pair exemplifying a morpheme structure condition while the other word violates this condition, judgment that the first word is more 'native-like' than the other is an indication of awareness of the condition involved. This appears to be a reasonable expectation.

The MSC's that Zimmer tested involved three vowel harmony restrictions, two of which replicate certain "very productive" vowel harmony phonological rules, while the third condition has no phonological counterpart. The results showed that the MSC's with phonological counterparts were productive, and the MSC with no phonological analog was unproductive. Zimmer offered the explanation that the productivity of the first two MSC's may have been the extended effect of the corresponding phonological rules, which he claimed were highly productive (in the sense of having many extant forms that obey the rules). Zimmer concluded that certain regularities linguists postulate, for example, the unproductive MSC, are not necessarily part of the speakers' grammar. Could the conclusion be generalized beyond this one MSC? Since it is relatively uninsightful to know that specific MSC's in specific languages are unreal, the point that Zimmer really wanted to make was that as long as there is at least one psychologically unreal MSC, then the theory of MSC's becomes suspect. By way of answering this question, two points need to be made clear.

First, the data do not unequivocally support the conclusion that the productivity of the two MSC's with phonological rule (hereafter, PR) duplicates is due to their status as PR's. If an MSC, which is duplicated by a formally productive PR, is found to be psychologically productive, it cannot be readily concluded that it is the phonological "ruleness" that induced productivity; phonological rule productivity is as much an issue as the productivity of morpheme structure conditions. The productivity of the first two MSC's could still be due to the fact that they were MSC's, and not to their status as PR's. (Another possibility is that the observed productivity was due to the status of the processes as both MSC's and PR's.) the point being made here is that the effect of one type of restriction has not been isolated from the effect of the other. The findings only allow the weak conclusion that MSC's not duplicated by phonological rules are unproductive, a conclusion that is not by any means unimportant, considering the current negative thinking on the necessity of MSC's. Secondly, since there appears to be an overlap in (underlying) MSC's and surface phonetic constraints (Shibatani, 1973), the possibility that the first two MSC's tested are in fact surface phonetic constraints would change the picture somewhat. Shibatani claims productivity for surface phonetic constraints, and thus it would have to be demonstrated that the productivity of the two Turkish MSC's tested is not due to the fact that they are (also) surface phonetic constraints. Zimmer's results only weakly suggest the unreality of MSC's.

But Zimmer raised an important question: If Turkish speakers do not make use of certain rules linguists formulate to describe base-internal vowel harmony in Turkish, how do they learn the correct forms? Zimmer suggested that "presumably they learn [the relevant] lexical items in their fully specified form and then simply repeat them" (p. 319-20). If it is indeed the case that speakers learn certain alternating forms by 'listing' (memorizing) their surface forms, this has grave consequences on a fundamental assumption of generative phonology. Hsieh

(1970) addressed himself to this question, a study that will be discussed shortly.

On the Psychological Manifestations of MSC's.

Ohala and Ohala (1975) reported two experiments on the psychological manifestations of morpheme structure constraints. Only the second experiment will be discussed, since it produced the less equivocal results, and also Because it illustrates a basic consideration in undertaking a phonological experiment.

The basic point is that, for any formal theory to be psychologically testable, it must make certain psychological claims, or, lacking such claims, it must at least justify certain psychological interpretations. Without at least the weak condition, it is obviously pointless to test for the psychological implications of a formal theory.

With this as a starting point, Ohala and Ohala interpreted Chomsky and Halle's notion of "degree of phonological admissibility" (Chomsky & Halle, 1968, p. 416) to presuppose the psychological claim that in judging the admissibility of a form that is not part of the speaker's lexicon, the speaker bases his judgment on only part of the form. The basis for this interpretation is Chomsky and Halle's suggestion to evaluate the closeness of a non-English form to English in terms of the segment (or combination of segments) that allows for the simplest "distinguishing rule" (the rule that distinguishes a form from a lexicon). For example, the forms "spøf" and "mløf" are judged equi-distant from English, since the focus of the simplest distinguishing rule in both cases is the non-English segment [ø]. The fact that the second word contains another violation of English MSC's--a nonpermissible initial cluster--is ignored in the Chomsky and Halle system. Thus, the Ohalas are justified in making the interpretation that for this aspect of Chomsky and Halle's system to have psychological relevance, speakers must use only that part of the form that the rule focuses on.

Greenberg and Jenkins (1964) proposed a "phoneme substitution procedure" to measure phonological distance. The Ohalas stated this procedure succintly as follows: "The degree of closeness of a given word to the native pattern would be inversely proportional to the number of zero-, one-, two-, and up to _n_-phoneme substitutions for the _n_ original phonemes in the word which resulted in existing words in the language" (p. 1). Illustrating the procedure would require more space than can be provided here. Suffice it to point out that this procedure takes into account the _whole_ form, in contrast to Chomsky and Halle's system which takes into account only part of a form. Greenberg and Jenkins, of course, intended their procedure to have a direct psychological interpretation, and in fact the results of their experiments showed that their procedure correlated very well with native speaker subjective judgment of lexical distance, a finding that has been corroborated in french (see D'Anglejan, Lambert, Tucker, & Greenberg, 1971). Do speakers make use of the whole form or only part of it in judging the form's closeness to the lexicon?

The Ohalas conducted an experiment to answer this question. Four pairs of nonsense forms were constructed such that the Chomsky and Halle model would predict that the words of each pair would be judged equi-distant to English, while the Greenberg and Jenkins procedure would predict that a member of each pair would be judged farther from English. The words were recorded on audio tape, and presented (as individual items, not as pairs) to the subjects, who rated the forms as to "closeness to English" on an 11-point scale. The results showed that "in all four word pairs, the difference in ratings is in the direction predicted by Greenberg and Jenkins and in three of these cases the difference is highly significant" (p. 6). The Ohalas concluded that the whole word and not just part of it contributes to speaker ratings of the word.

What is at stake here is the answer to the one question that has perplexed researchers in this area of linguistics: Do subjects employ rules of the kind proposed in formal linguistics? The Chomsky and Halle system presupposes that subjects advert to rules that are independent of specific elements in the lexicon. The Greenberg and Jenkins procedure presupposes that specific items in the lexicon, not general rules, are consulted; and hence, that, at least in this particular area of the grammar, generative rules are not necessary. The Ohalas wrote: "We believe the results of both studies also strongly imply (although perhaps do not conclusively prove) that the speakers make reference to the words in their lexicon, not to lexicon-independent rules, to make judgements in tests such as these" (p. 6). If such a conclusion can be maintained, it constitutes a revealing finding on the interaction between mental and formal grammars.

It appears, however, that at least one aspect of the experiment is open to question. In this kind of an experiment, the essential question is whether the subjects judged "mløf" to be farther from English than "spøf" on the basis of (a) the non-English segment [ø], (b) on the impermissible cluster [ml], or (c) on both. Judgments based on (a) and (b) take into consideration only parts of a form; judgments based on (c) involve the whole form (ignoring the presumably irrelevant last segment). It is safe to say that subjects did not base their judgment on the non-English segment alone, otherwise both words would have received equal rating. The finding that "mløf" was judged to be farther from English, however, eliminates neither (b) nor (c). The second alternative does not support the conclusion made by the Ohalas. If subjects based their judgment of "mløf" on the nonpermissible cluster alone, and if a nonpermissible cluster of this type happens to have been judged to be more "un-English" than the non-English segment [ø], then "mløf" would be judged farther from English. This is a logical possibility that has not been eliminated. If this was what happened, then only part of the word was being consulted, and this would mean that the Chomsky and Halle prediction can be maintained, but that, in these specific cases, they were wrong in identifying that part of the form that would serve as focus in processing; in other words, that a Chomsky-Halle distinguishing rule focusing on [ø] would be wrong for just these words.

An Experiment on an Abstract Segment

M. Ohala (1974) attempted to test directly for the reality of an abstract segment, using Hindi as the language for consideration. In Hindi, certain forms with a medial cluster alternate with forms in which this cluster is broken by schwa, for example, *p'Isəl*/*p'Isla* 'slip/slipped'. Ohala suggested postulating for such words an underlying form with schwa, and to derive the forms with the medial cluster by a schwa-deletion rule. Ohala further observed that there are a few noun morphemes, which are pronounced with a cluster at the phonetic level, but which have no alternating forms with schwa; for example, *g'o:sla*. Ohala then asked whether schwa should be posited in the underlying forms of these morphemes. Observing that the suffix -*iya* blocks schwa-deletion, Ohala wondered if schwa would "re-appear" in *g'o:sla* words in suffixation with *-iya* (these words do not normally take this suffix). If the resulting forms were pronounced with schwa, this was to be considered evidence for an underlying abstract segment. A word-formation test was conducted, and on the basis of the results of this test, Ohala concluded that for some speakers, the existence of schwa in the underlying forms of these morphemes is psychologically real. Before the results are discussed, several points need to be clarified.

First, is there any descriptive necessity for postulating an underlying schwa in *g'o:sla*-type words? Ohala observed that they have no alternating forms with schwa; if this is the case, then there are no alternations to be accounted for. In contrast, *p'Isəl* alternates with *p'Isla*. Moreover, there appears to be no phonological process that depends on the existence of underlying schwa in these words. In contrast, consider that the underlying abstract vowels postulated for Nupe, discussed in the first chapter, are at least necessary to motivate the surface occurrences of certain palatalized and labialized consonants. Could there be a morpheme structure condition against underlying clusters in Hindi? Thus, it seems that the descriptive basis for postulating schwa in this small set of words has not been made clear.

Second, there is the question of whether schwa is an abstract segment in Hindi. Schwa appears in surface phonetics in other alternations in Hindi. Consider that the abstract segments that have caused considerable discussion do not appear in the surface phonetics of the languages for which they have been postulated, for example, Chomsky and Halle's velar spirant for English, Hyman's low vowels for Nupe, Kisseberth's (1969) long *u* for Yawelmani, Sanskrit short *e* (discussed in Kiparsky, 1973a), Brame's (1972) voiced pharyngeal spirant for Maltese. Furthermore, let us suppose that, following Ohala, the fact that *g'o:sla* words do not take the suffix -*iya* is merely a morphological accident. Had the accident not occurred, Hindi would have alternations of the sort *g'o:sla*/*g'o:səliya*. Presumably the lexical representation of the stem is something like /*g'o:səl*(*a*)/. This is the sort of representation that would be postulated if the experiment yields positive results. But this representation, given the postulated alternation, is not an abstract representation, i.e., not "imaginary" in the sense of Crothers (1971), but only a "mixed" representation, since schwa would appear in the putative variant *g'o:səliya*. In other words, the absence of a variant of these words with schwa is only a morphological accident. And since, as Crothers noted, the abstractness controversy arose over imaginary segments, the relevance to the abstractness controversy of the Ohala experiment would vanish if schwa is in fact not abstract in Hindi.

A third point is that there may be no intrinsic objection to abstract segments at all. The spirit of

Kiparsky's objection, as I understand it, is against absolute neutralization rules, and only indirectly against abstract segments, since, at that time, all solutions to abstract segments that have been proposed made use of absolute neutralization. But there can be non-neutralization solutions to abstract segments, for example, Krohn's (1976) feature-sequencing solution to the Nupe problem. In fact, the objection may be against any type of context-free phonological rules, neutralizing or not; included in this set is the non-neutralizing, context-free rule that derives the diphthong in the word "boy" from an underlying /œ/, which is its only source, a rule proposed in Chomsky and Halle (1968). If experimental evidence forces us to postulate schwa in the lexical representation of *g'o:sla* words, one does not need an absolute neutralization rule to get rid of the schwa in surface; one only needs to extend the coverage of a well-motivated schwa-deletion rule to include this small set of words.

In summary, then, there may be no descriptive necessity for postulating underlying schwa in *g'o:sla* words. Moreover, even if experimental evidence favors such a representation, schwa may not be an abstract segment in Hindi. Furthermore, the bone of contention may not be abstract segments but absolute neutralization rules, a type of rule that will not be resorted to to insure the non-appearance in surface phonetics of the segment schwa.

Thus, if the points raised here represent correct assessment of the facts, much of the theoretical significance of Ohala's study would diminish.

The results of Ohala's experiment showed that of the 70% consistent responders to five *g'o:sla*-type stimulus words, 49% responded with no schwa, and only 21% responded with a schwa. In other words, less than one-third of the consistent responders manifested a type of behavior that supports the psychological reality of the 'abstract' schwa. This is not an impressive number, but, following Ohala, it must be admitted that these schwa responses cannot be denied, either. Ohala, therefore, concluded that at least for these speakers, this class of word has to be represented with underlying schwa. This conclusion raises the interesting question of whether the psychological reality of linguistic constructs is something to be decided on a grammar by grammar basis.

This writer agrees completely with Ohala on the pressing need for *any* external evidence that may shed some light on the abstractness controversy. This question, however, is probably the most difficult to handle experimentally. The very nature of the objects under study--abstract segments--poses an almost insurmountable barrier to direct experimentation. By the strong definition of abstract segments here imposed, abstract segments are not part of the surface sounds of the language. Processing of linguistic materials is an active interaction between present knowledge of linguistic materials and the nature of the materials to be processed. How, then, can abstract sounds be presented to the subjects in such a way that the experimenter is certain that the subjects perceive them as such and not as some reinterpreted categories? Fromkin (1971) expressed doubt that underlying strings in general can be tested for reality: "But behavioral tests *of the sort* which reveals the *blick* knowledge would fail in the case of admissible *underlying* sequences.... There seem to be no 'external criteria' by which we can decide on the 'reality' of the underlying strings" (pp. 78-9).

The Reality of an Intermediate Representation

Generative rules describe processes of sound change, whereas taxonomic rules express equivalence relations among sounds. A generative rule is of the form "A becomes/changes to B in the context X", whereas a taxonomic rule expresses the observation that "A is B in the context X". Because it is process-oriented, the generative system allows for intermediate stages between the initial state and the final state. Is there any reality to these intermediate stages? If it can be shown that derivational intermediate stages that generative rules require have mental reality, then this may be viewed as a piece of evidence in support of the generative system. Anisfeld (1969) addressed this question.

English has a small set of verb-adjective pairs that exhibit the alternations t,d/s, for example, permit-permissive, submit-submissive, extend-extensive, decide-decisive. The t/s alternation appears to be limited to words of Latin origin whose root is "mit", and only when in suffixation with the adjectival suffix "-ive" (cf. permi[sh]ion). In a generative description of these facts (see Chomsky & Halle, 1968, p. 229), a rule is proposed that changes the dentals /t d/ to the spirants /s z/, respectively. This rule accounts for the t/s alternation but not for the other alternation. An additional rule is required to change /z/ to /s/ to produce extensive; thus, exten[z]ive serves as an intermediate stage in the derivation of extensive.

Synthetic verbs were constructed (e.g., flamit, garlude), and each verb was paired with "-ive"-adjective forms whereby the crucial sound was alternately one of the following: [s], [sh], [th], [f]. Subjects were asked to judge the acceptability of the "-ive"-forms as adjective counterparts of the verbs. Anisfeld predicted that the pair garlude-garluzive would receive significantly better acceptability ratings than the pair flamit-flamizive. The results showed that this was just the case.

But why should such a demonstration constitute psychological evidence for the reality of the intermediate "-zive" stage in the extend-extensive derivation?

Apparently such a conclusion is based on the following assumptions: First, acceptability judgments of the sort elicited in the Anisfeld experiment are based on the adjective forms as intermediate forms; and second, any form which is part of a derivational history is judged to be a more acceptable derived from than a form which is not part of the derivational history.[1] The second assumption is an empirical question; there is no a priori reason to suppose that this should be the case. Data relevant to the question could have been easily gathered had the sound [s] been included in the set of crucial sounds. Thus, if the second assumption is valid, it is to be expected that flamisive is judged more acceptable than flamizive since the latter form is not part of the derivation.

The first assumption is more properly a hypothesis to be tested than an assumption. The basis on which acceptability judgments are made is difficult to ascertain. In the Anisfeld experiment, judgment could have been made on the basis of 'phonemic distance'. It will be noticed that d --> z requires one change, which is spirantization, whereas t --> z requires two changes -- spirantization and voicing. Moreover, Lackner and Goldstein (1975) have 'verbal transformation' data that show that auditory misperceptions with respect to the feature voicing is uniderectional, that is, voiced to voiceless. Thus, the greater distance of t --> z (yermit-yermizive) and the uniderectional voicing effect observed by Lackner and Goldstein would predict more favorable ratings for the pair that exemplifies d --> z (garlude-garluzive), which was precisely the pattern of results obtained. In other words, the results could be explained in terms unrelated to the issue of the reality of intermediate representations.

---

[1] A more basic assumption is that stages in a derivation are accessible and are being accessed to in processing. This is an assumption that is hardly open to question, unless the existence of any level of representation other than the surface level is to be denied. A hypothesis can be made here, namely, that all representations or derivational stages have equal reality. This hypothesis predicts that the surface garlusive will be judged no more acceptable than the intermediate form garluzive. A modified hypothesis is that reality of representations is directly proportional to shallowness of representation.

Testing for the Reality of Phonological Rules

On the Generativity of Phonological Rules

Hsieh's (1970) experiment dealt with the productivity of a set of phonological rules--tone sandhi rules in Taiwanese. This study, however, is more than a rule productivity study; Hsieh questioned the generativity of phonological rules: "Is there any psychological reality to the claim that speakers consciously or unconsciously apply a set of phonological rules to some abstract 'underlying forms' to get the 'surface forms'?" (p. 489).

Hsieh constructed artificial but permissible morphemes to fill in certain gaps in the lexicon of Taiwanese. These artificial morphemes were given meaning, and put in a paradigm productive for tone sandhi, i.e., as verb or as object in the structure #one-syllable verb+one-syllable object#. Hsieh then hypothesized that if the tone alternations were rule-governed, then tone sandhi would create the appropriate tone patterns for these artificial morphemes. An experiment was conducted to test for this expectation. Sandhi phrases were constructed as described above, and subjects were asked to read the phrases aloud.

The results showed that tone sandhi was productive only when the made-up items occupied the object position, i.e., before #, but not when they occupied the verb position, i.e., in the environment before +. This seems to indicate, Hsieh concluded, that for Taiwanese speakers, only the second part of the tone sandhi rule is productive. Accidental gaps seem to be more sensitive to the 'before #' part of the rule than to the 'before +' part. Hsieh remarked that generative phonology cannot provide a satisfactory explanation for this result. He then outlined a theory of phonology that, he believed, would provide a satisfactory explanation. This theory has the following properties:

P1. In the lexicon, for each morpheme, all the surface forms are listed and the exact occurrence of each variation form is stated.

P2. When the speaker wants to utter a particular morpheme in an utterance he consults the lexicon and the lexicon will tell him which surface form of the morpheme to choose.

> P3. Any regularity obtained in the surface variations may become a part of the knowledge of a speaker, but it is not necessarily so. In case the regularity does become a part of the knowledge of a speaker, such regularity is stated as a phonological rule (in the generative sense). (p. 497)

Thus, the morphophonemic alternations for the morpheme be 53 'to buy' are to be stored in the lexicon as be 53 (citation form), be 53 (before #), and be 55 (before +). Hsieh then suggested that there is a "phonological rule" that states that the citation form is the same as the "before #" form. Thus, in the case of the artificially created pai 55, it is stored as such as the citation form, and a phonological rule guarantees that the tone 55 is retained in the object (before #) position. If it were used as a verb (i.e., before +), subjects would not know its derived tone since there is no corresponding representation in the lexicon.

The property of this theory of phonology that directly provides an explanation for Hsieh's results is P1, which specifies that morpheme alternants as well as their environments are listed in the lexicon. P1, in conjunction with P2 (which is supplemented by a set of distributional, not generative, phonological rules), denies the generativity of a lexical grammar. The one unclear aspect of Hsieh's suggestion is P3. Given that the first two properties are sufficient to produce the correct forms in all cases, what are generative rules for? The first part of P3 implies that a list lexicon, not being a description of regularities but merely a list of forms, is not part of speaker knowledge, a point that is difficult to accept. Furthermore, if a surface generalization is abstracted and expressed as a generative rule, as P3 requires, the question arises as to whether certain of the surface alternants, specifically the derived forms, are then removed from the lexicon.

If one denies the relevance of generative rules in the lexicon, one has to explain how lexical items are related, and furthermore, how a new morpheme ("new" to the individual's lexicon) is eventually decided upon as a member of a class of morphemes which behave similarly with respect to a sound change. Hsieh (1975, 1976) suggested that speakers analogize with known forms. Hsieh suggested that speakers use analogy in handling unfamiliar words, including nonsense words used in linguistic experiments. Thus, Hsieh claimed that in productivity tests, occasional item-particular successes are due to isolated cases of automatic or obvious analogy; and that, in the majority of cases, forms are merely repeated. Hsieh interpreted the negative results to mean that the forms were simply memorized or mechanically listed in the lexicon.

Used as a strategy for language learning, analogy is not a perfect, efficient machine. The subject may be too slow to come up with the correct analysis, or he may make the wrong analysis, or the "rule" may be too advanced for his linguistic development. The imperfections of analogy thus result in rule opacity; rules do not apply when they should, or they apply seemingly randomly, or the learner applies early approximations of the rules, thus producing "temporary" forms. All these, Hsieh suggested, result in a lexicon that, from the point of view of the "ideal" speaker, is imperfect. Hsieh suggested that the results he obtained in these last two experiments support, and are in turn explained by, Wang's (1969) principle of lexical diffusion, which stipulates that while phonetic changes may be abrupt, lexical changes are gradual, thus "stranding", so to speak, many existing forms. These forms may or may not be regularized as the changes get diffused across the lexicon. Wang's principle has been formulated primarily to account for diachronic changes, and Hsieh's studies seem to provide evidence that this process is synchronically recapitulated (see also Hsieh, 1972, for an attempt to apply this concept in explaining certain child language acquisition data).

Experiments on the English Stress Rules

One of the most complex sets of phonological rules ever written for any language is the set of stress-assignment rules Chomsky and Halle (1968) have proposed for English. The rules assign primary stress on basic morphemes, 'copy' this stress onto certain syllables under certain conditions, and adjust previously assigned ones. The rules apply at every morpheme derivational stage, i.e., cyclically. Some of the rules conflate a number of subrules that are conjunctively or disjunctively ordered, necessitating the employment of parentheses, braces, angled brackets, and lexical category subscripts. The rules figure prominently in the attempt by Chomsky and Halle to motivate the cycle.

The rules do not always give the correct results. Ross (1972) suggested revisions; he succeeded in accounting for more data, but he also complicated the rules further. Other revisions were proposed in Halle and Keyser (1971) and Halle (1973). Proposals have been made to 'de-cycle' the application of the rules, for example, Schane (1974), Settera (1974), and Sloat (1974). However, notwithstanding the many proposals on how to account for stress in English, one assumption is generally shared, namely that stress is predictable from the phonetic properties of words. Halle and Keyser (1971) consider stress predictability as "one of the most surprising discoveries of Chomsky and Halle (1968)" (p. 3). If phonetic properties (and syllable structure) completely determine stress placement, (but cf. Schmerling, 1971 for evidence that the stress of certain noun compounds must be represented in the lexicon), native-like nonsense words can be very useful in testing for the predictions of the rules. Three such studies are discussed below.

Ladefoged and Fromkin (1968). Twelve nonsense words were made up, for example, sitrenide, and put in sentence frames, thus: 'He is going to sitrenide the paper'.

Twenty-five subjects, all linguistics students, were asked to transcribe how they would pronounce the words. Except for two words, the results showed that the subjects were in agreement with Chomsky and Halle's formulation. The data, however, may not be reliable; spelling effects and analogical patterning have not been controlled for. Furthermore, the fact that a subject transcribes a vowel sound as tense and stressed does not necessarily mean that the vowel has been perceived as underlyingly tense and that, *on this basis*, stress has been assigned on it. The possibility that the vowel has been made tense *because* it has been assigned stress has not been ruled out. The Chomsky-Halle rules presuppose an implicit ordering of tensing preceding stress assignment. Subjects' performance, however, merely suggests that at the level of pronunciation, stress and tensing occur together.

Jackson (*1972*). Unnested disyllabic words with a tense vowel in the final syllable have either final stress (with stress contour 0-1, for example, *a^{0}nti^{1}que*, *ca^{0}noe^1*, *ca^{0}pri^{1}ce*, etc.), or they may have the stress contour 1-3, for example, *a^{1}rgy^{3}le*, *mi^{1}cro^{3}be*, *a^{1}mpe^{3}re*, etc. (In this section, the superscripts refer to stress levels: 0 is unstressed, 1 is primary stress, and 3 is tertiary.) In this second set of words, stress seems to have been retracted (with accompanying lowering of final stress to tertiary). As Halle and Keyser (1971, p. 62) noted, it is easy to revise the relevant rule to cover these cases of retraction, but they refrained from doing so on the ground that this will result in a weakening of the grammar (for their arguments, see pp. 62-3). They decided instead to retain the rule as formulated, and, on the basis of a larger inventory of 0-1 words, suggested to characterize the 0-1 words as 'regular', and the 1-3 words as 'irregular' and are thus treated as exceptions to the rule. Jackson wanted to find out if native speakers agree with Halle and Keyser in treating 0-1 words as regular and 1-3 words as irregular.

Forty-eight nonsense words were presented in sentence contexts to 34 subjects. Effort was made to make these words similar to the words in the Halle and Keyser lists of 'regular' and 'irregular' words, and also to have the orthography suggest the presence of a tense underlying final vowel, for example, pelcoze, garfile, arpete, hestoy, etc. Results showed that 75% of the nonsense words used as nouns, 64% of the nonsense forms used as verbs, and 75% of the nonsense forms used as adjectives received non-predicted stress. These percentages of items were assigned the irregular 1-3 pattern. The conclusion Jackson made was that what the grammarians Halle and Keyser believed to be the irregular stress contour turned out to be, from the point of view of non-linguist speakers, the regular pattern.

In this experiment, it is crucial that the final vowel segment be perceived by the subjects as tense, so that primary stress may be assigned to it. Jackson made an effort to achieve this by representing the last syllable in orthographic form that suggests vowel tenseness. There is, however, no guarantee nor indication that such a procedure had succeeded. It is unresolved whether or not the subjects did indeed perceive the last syllabic as tense.

Nessly (*1973a*,*b*). Nessly subjected the Chomsky-Halle English stress rules to experimentation. In the first paper, a somewhat unexpected but interesting result is the observation that subjects tended to divide words into morpheme-like subconstituents, and to assign stress on these 'morphemes' in terms of simpler stress rules. For example, *balderdash* is broken down to *balder* and *dash*. This is an attractive possibility that must be pursued. Simpler primary rules may be easier to formulate if the domain of application is no more than two syllables. All that will then be needed is a small set of secondary rules that will take care of the adjustments required when such syllable groups are concatenated to form a word. For example, *ba^{1}lder* is assigned stress on the first syllable as a consequence of the regular main stress rule; *da^{1}sh*, being monosyllabic, gets primary stress. Then an adjustment rule lowers the primary stress on *dash* to tertiary when concatenated with a word which, like *balder*, has a primary stress, giving the correct *ba^{1}lderda^{3}sh*.

In the second paper, Nessly dealt with procedural aspects of experimentation, of the sort discussed earlier.

On the question of regular or standard and irregular pronunciations, he has evidence that linguists' pronouncements in this regard are usually not borne out by experimental data. For example, Ross (1972) considers the penultimate stress on *disa'ster* to be irregular (his rules predict the incorrect **di'saster*; cf. *si'nister*). Nessly asked his subjects to pronounce the novel form *capaster* (on analogy with *disaster*; this, of course, is the weak point in Nessly's procedure). His reasoning is that if the penultimate stress on *disa'ster* is exceptional, as claimed by Ross, and if what is regular is antepenultimate stress, then the subjects should put antepenultimate stress on the test word, thus *ca'paster*. However, the subjects put stress on the penult, *capa'ster*, suggesting that, contrary to Ross' suggestion, the penultimate stress on *disa'ster* is regular. As another example, Ross considers the initial stress in *ca'lumny* to be regular. Nessly asked his subjects to pronounce a similar novel form, *metumny*. Eleven of eleven subjects put the stress on the penult, *metu'mny*, again contrary to the prediction, suggesting that what Ross considered to be regular native speakers treated as exceptional.

Even standard dictionary pronunciations may not be so standard after all. The Oxford English Dictionary (1933) and Kenyon and Knott (1953) put penultimate stress on *vaga'ry*. Random House (1966) and Merriam-Webster's (1966) list the word both as *va'gary* and *vaga'ry*. One expects, as

Nessly does, that the two variants would split the subjects, or, that the variant vaga'ry, since it is the sole entry in two sources, would be favored. Ten of Nessly's 12 subjects favored va'gary, suggesting the irregularity of what the OED and Kenyon and Knott consider to be regular pronunciation. As another example, Nessly noted that patina is cited in the OED, Kenyon and Knott, Jones, and Random House only as pa'tina, while Merriam-Webster's has both pa'tina and pati'na. The expectation is that pa'tina should be favored. Nessly's results showed that thirteen of his thirteen subjects responded only with pati'na, suggesting again the exceptionality of a regular dictionary entry.

The point here is that these 'standard' dictionary entries often serve as bases for linguistic analyses, and if native speakers do not agree with them, it is obvious that little faith can be placed in these analyses. Furthermore, what linguists may consider to be regular or irregular pronunciation, decided mainly on the basis of the number of forms exemplifying a paradigm, may not receive supporting opinion from native speakers. The fact that a stress pattern is listed as standard, or may have many exemplars, does not insure its productivity. Subjects may not use it for the reason that, Nessly suggested, it may be less preferred, or -- and this is what is relevant to rule construction -- it is unstable and is in the process of being eliminated, leaving behind its trace in many words.

J. Ohala (1974b). The experiments of J. Ohala must be mentioned here. It is unfortunate that, because of the tentative nature of the report released so far, the study cannot be given the full attention that it deserves. In the absence of a fuller description of the experiments, the findings are merely reproduced below; they will not be discussed.

Ohala tested for the productivity of several of Chomsky and Halle's rules. The rules and the results are as follows: 1) vowel laxing, e.g., [ey] becomes [æ], as in "profane/profanity" (result: unproductive); 2) vowel tensing, e.g., [æ] becomes [ey], as in 'manic/mania' (result: highly productive); 3) the vowel tensing rule that tenses a vowel when followed by another vowel, as in 'algebra/algebraic' (result: little evidence of productivity); 4) a stress rule that shifts stress to the syllable before certain suffixes, for example, the suffix -ian, thus, 'muˈsic/musiˈcian' (result: highly productive); 5) s-voicing, for example, 'Paris/Parisian' (result: unproductive), and 6) velar softening, e.g. [k] becomes [s], as in 'public/publicity' (result: marginally productive).

Two Experiments on the Vowel Alternation Rules

The studies discussed in this section, Steinberg and Krohn (1975) and Moskowitz (1973), tested for speaker knowledge of vowel alternation in English. These studies reported conflicting results. Steinberg and Krohn provided disconfirmatory evidence and Moskowitz supplied positive evidence. Methodological considerations that quite possibly have influenced the divergent results are discussed below.

Steinberg and Krohn collected real English bases which when combined with any of the suffixes -*ic*, -*ical*, -*ity*, -*ify*, -*ish* produce "meaningful" but nonexistent forms, for example, *snide* + *ity* --> "snidity". The experimental procedure supplied a verbal context for a particular base and suffix combination, and the subjects, who were university students, were asked to supply the combination. Steinberg and Krohn assumed that a response of the type *sn[I]dity* is supportive of VAR; whereas the response *sn[ay]dity* (i.e., no change in the quality of the crucial vowel) is to be considered disconfirmatory. Nearly 90 per cent of the valid responses were of the disconfirmatory type. Since the critical vowel sounds remained unchanged in the derived forms, Steinberg and Krohn concluded that the subjects did not "know" the vowel shift. On the basis of these results, Steinberg and Krohn questioned the validity of the vowel shift rule.

Moskowitz, on the other hand, used made-up forms as bases. To elicit responses, the concept-formation technique was employed. The subjects were told that to made-up words they were to add -ity, and that they were to pronounce the resulting words. One of the nonsense words used was [fayp]. A response exemplifying the 'predicted' alternate form, which, in this case is [fIpItiy], was acknowledged to be correct; any other responses were commented on as erroneous and the expected response was supplied. The concept to be 'formed' was the correct vowel alternant. Two control conditions were set up, both of which exemplified alternations not in consonance with VAR. The subjects were children aged 9 to 12 years. The results showed that subjects who did the correct alternation conditions performed to criterion, whereas only two of the 20 subjects who belonged to the incorrect alternations conditions learned the alternations. In other words, alternants predicted by VAR were more easily discoverable, or learnable. On this basis, Moskowitz concluded that vowel shift is part of the subjects' linguistic knowledge.

Why should such diametrically opposed results be obtained? The choice of materials may have contributed a great deal to the negative results of Steinberg and Krohn. They used real English words as base forms of the stimuli, whereas Moskowitz used nonsense bases. As pointed out earlier, the use of made-up forms in productivity tests is motivated by the consideration that with real words, it is uncertain whether the rule or phenomenon being tested has mediated a correct form, or whether a correct response is simply a memorized form. With made-up stimuli, the experimenter is reasonably certain that correct forms have not been memorized. In the Steinberg and Krohn study, while a base plus affix combination resulted in a derived form that was a non-existing English item, the *crucial* part--the part that carries the critical vowel sound--was a real English word. Therefore, as far as the relevant materials were concerned, they were not novel to the subjects. The subjects assigned to these bases their regular pronunciations. Thus, the results obtained may have been due to the fact that the pressure of the regular pronunciation of the real base words was much too strong to allow for the expression of vowel shift (if the process is at all real).

The near unanimity of judgment by Steinberg and Krohn's subjects (90% of the valid responses were "nonshifted" forms) may have been an artifact of the experimental task. The task required the subjects to express their preference between two derived words, for example, *maze* + *ic* and *maze* + *ity*. Since both suffixes supposedly trigger vowel shift, the actual choice was of course irrelevant to the outcome of the experiment. The relevant point was how the base vowel would be pronounced. Since the subject was told to add the suffix to the real base, this was what was invariably done. Suppose the subject replied with *mazic*, with no vowel change. The experimenter could not but accept this response. Such an acceptance, in effect, was an acknowledgement that the response was correct. The subject would then proceed to simply add suffixes to the base words with no concomitant vowel shift, thinking that this was all there was to it. In other words, the procedure had no adequate control for response bias. Thus, on the whole, Steinberg and Krohn's experimental task may not have been sensitive enough to allow for the expression of vowel shift.

The technique of concept-formation, which Moskowitz employed, has a built-in control for response bias. Subjects who formed words by simply adding the suffix to the base received negative feedback, and, hence, were forced to search for other strategies. Thus, subjects who belonged to the correct alternations conditions "found" the alternation generalization and used it, but subjects who belonged to the incorrect alternations conditions had no comparable generalization available to them, hence, they did not learn these incorrect alternations.

Moskowitz' study is interesting in another respect: It attempted to test for the appropriateness of certain aspects of the form of Chomsky and Halle's shift rule. The Chomsky-Halle formulation covered three front vowel alternations and three back vowel alternations. Should all these alternations be expressed as one generalization? In an attempt to answer this question, Moskowitz designed her experiment such that the first part (Part A) of the test was limited to two front vowel alternations. Then, after criterion (which was set at ten correct responses in a row), or, if criterion was not reached, when all 72 stimuli had been presented, nonsense exemplars of the third front vowel alternation and of the back vowel alternations were presented immediately and without additional instruction. Moskowitz hypothesized that correct production of these other items would Indicate that

> ...criterial performance on A [the first two front vowel alternations] resulted from prior knowledge which encompassed the three vowel pairs, while incorrect production...following criterial performance on A would indicate that A had involved only the learning of two specific vowel alternations. (p. 234)

Criterial performance on these other alternations was to be taken as evidence for one unified psychologically real vowel alternation generalization. Responses on the first item exemplifying the third front vowel alternation and the back vowel alternations should, therefore, be revealing. Results showed that of the eight subjects who belonged to the correct alternations conditions, only one did not give the correct alternant of the third front vowel alternation. Half of the subjects correctly identified all six items exemplifying the third front vowel alternation; the average correct response was 4.6 of a possible 6, indicating fairly successful overall performance. Performance on the back vowel alternations, however, was very poor for all subjects.

Moskowitz put to test one other formal aspect of the

Chomsky-Halle shift rule. The rule was formulated independently of vowel tensing and diphthongization. To test for the independence of vowel shift from tensing and diphthongization, Moskowitz had some of her subjects work on pairs which illustrated tensing and diphthongization but not shift, e.g., [tiyb-tIbItiy]. Another group of subjects were given pairs exemplifying tensing, diphthongization, and shift, but the shift was incorrect for English. Thus, the nonshifting alternations were formally simpler in that they required one less rule. The results showed that such nonshifting alternations were hardest to learn, even harder than alternations with incorrect shift. It would seem that any shift, even an incorrect one, was better than no shift at all. Moskowitz concluded that shift is not separable from tensing and diphthongization. This finding, and the finding that while criterial performance on two front vowel alternations triggered correct performance on another front vowel alternation but not on any back vowel alternations, have lead Moskowitz to suggest that "the SPE proposal about vowel shift has no resemblance to the functional grammar of native speakers" (p. 249).

Since the present study is also on vowel shift, it should be made clear why it was believed necessary to do another study using this aspect of English phonology as the testing ground. The more obvious reason relates to replicability of experimental results. The reliability of experimental results is supported if the results are replicated. The results of the studies just discussed not only fail to support each other but are in fact contradictory. Additional data may shed some light on the issue. Furthermore, Moskowitz reported that vowel shift, contrary to expectations, did not exhibit any inhibitory effects. The present writer believes that interference did not occur because the task was too easy; it is hoped that the proactive inhibitory effect of vowel alternation, which is a strong indication that the generalization is operating, will be observed in the present study.

Summary, and Some Suggestions on Procedure

In the first half of this chapter, an attempt was made to substantiate the basic methodological assumption of the present study, which is the pervasive effect of natural language habits on performance in verbal learning. The studies discussed showed that stimulus materials were re-organized in perception and in learning according to certain phonological 'habits' of the subjects. In the second half, rule-validation experiments were reviewed; while these studies raised very important questions, a number of them suffer from loose methodology that resulted in uninterpretable results. The following are some suggestions on how some of these procedural problems may be avoided.

The stimulus materials should be novel forms to ward off familiarity effects. Though novel, the materials should look like real English words so that English phonological rules apply on them, but they should not resemble real English words to a degree that they remind subjects of particular English words; otherwise analogy with these real words may occur.

Analogizing from earlier responses can be prevented by judicious arrangement of materials, by randomization of materials across subjects, or by interspersing them with distractors. If none of these is possible or desirable, a task that prevents the use of such a strategy should be employed. Direct response modelling is relatively difficult to do in a learning task.

If the rule to be tested is particular to words of a certain historical origin, the stimulus materials should be made to look as if they belong to that class of words. If one is testing for vowel shift, for example, which, according to Chomsky and Halle (1968, p. 178) is found primarily (but not solely) in words of Romance origin, the use of Romance affixes may be resorted to. For example, the made-up word surkete, because of the Latin prefix sur-, looks more Romance than Germanic.

Spelling effects can, of course, be minimized by presenting materials aurally. There is still the possibility, suggested by Ohala (1974b, p. 41) that responses are mediated by "orthographic images" subjects form of aurally presented materials.

The Experiment

The purpose of the present study is to investigate the psychological relevance of the generalizations underlying the Vowel Alternation Rules. A generalization is considered psychologically relevant if it has demonstrable effects on certain psychological operations, for example, on learning. An aural paired-associate learning task was administered to test for the learnability of five of the six major vowel alternations discussed in Chapter 1.

In paired-associate learning, subjects generally impose a structure on nonsense materials (Glanzer 1967); whenever possible, they make use of "natural language mediators" (Prytulak 1971), which are established as secondary stimuli. It now appears that natural language mediation is not restricted to the use of word forms alone. Links of a syntactic nature may be established between stimulus and response. Glanzer (1962) claimed that analyzability of the stimulus and response items as a syntactic unit facilitates learning. Baker and Prideaux (1975) proposed to view this kind of mediation as 'syntactic facilitation'. Since the theoretical status of a syntactic generalization, which induces syntactic facilitation, is no different from a phonological generalization, it is expected that phonological knowledge should likewise facilitate learning.

The basic assumption of the present experiment is as follows: If a linguistic generalization has a psychological basis, it will guide linguistic performance in measurable ways. Or, complementarily, if a linguistic generalization has a psychological basis, it will interfere in the learning of materials that violate the generalization. Hence, if the alternation generalizations which form the basis of the Vowel Alternation Rules are real to the speakers, they will provide internal guidance in tasks involving the use of vowel alternations. In a learning task, for example, the availability of such a generalization should result in facilitated or retarded learning depending on the positive or negative nature of the materials.

On the basis of this assumption, the experimental hypothesis is stated as follows:

> Subjects will find it easier to learn alternations that have real exemplars in their language than alternations that do not have real exemplars in their language.

Thus, real English vowel alternations should be found easier to learn by native speakers of English than any other pairings of base and derived vowels, which would result in alternations that have no exemplars in English. The manner in which a psychologically real vowel alternation generalization will facilitate the learning of real alternations in a paired-associate learning situation will be explained.

Consider a three-stage model of paired-associate learning (McGuire, 1961). First the stimuli are discriminated from one another. Then the responses are learned, a process that is essentially of the recall type. Finally, the stimuli are associated with the appropriate responses. If now a vowel alternation generalization of the sort linguists postulate is part of the speaker's mental grammar, this should provide a link between stimulus and response. The ready availability of such a link simplifies the associative phase, and, in this sense, improvements in learning should result. A second way by which an available mental generalization should facilitate learning is in the minimization of response learning, or, in Bower's (1970) term, in the restriction of available responses. If indeed a mental vowel alternation generalization exists, which specifies what base vowels are paired with what derived vowels, speakers must have 'knowledge' of the vowel pairings. Hence, they do not have to learn the derived vowels that crucially distinguish the response terms. In fact, given a stimulus term, the response term is completely recoverable. The subject needs only to add the suffix -ity to the stimulus term, and, using the generalization, change the shape of the crucial vowel. Paired-associate learning of real alternations is thereby reduced to stimulus learning. Such a simplification of the learning process is not suggested by alternations that are not part of the language of the subject. The expectation in these cases is that these non-existing vowel alternations will be harder to learn. The remainder of this study is concerned with the experiment conducted to test for these expectations.

CHAPTER THREE

METHOD

Subjects

The subjects in both the acquisition and the transfer experiments were 40 native speakers of Canadian English. Thirty subjects were fulfilling a service requirement for an introductory psychology course at the University of Alberta. The remaining 10 subjects were students in an introductory course in linguistics. None of the *S*s reported having any hearing impairment.

Experiment 1: Acquisition

Materials

The materials were pairs of nonsense words, for example, *subk[ay]pe*-*subk[I]pity*. Following standard linguistic practice, the shorter form will be referred to as the *base* term and the longer form as the corresponding *derived* term. A base-derived pair has the same 'root' form; the forms differ in the shape of the crucial vowel and in the fact that the derived word ends in -*ity*. The base forms are disyllabic, with the first syllable being one of the following prefixes of Romance origin: *sur*-, *mal*-, *de*-, *pre*-, *per*-, *sub*-, *im*-, *ex*-, *mis*-, *en*-, *dis*-, *pur*-. The use of Romance prefixes was intended to make the materials look like words of Romance origin. The second syllable has the structure CVC, where C is one of the consonants p, t, k, b, d, g, f, h, m, n, r, l, except that h, r, and l never appear after the vowel.

Base words were made up in the following manner. A consonant was randomly paired with another consonant, and the relevant vowel was inserted. If the resulting form turned out to be in Thorndike and Lorge's (1944) list of 1,000 most common words, or if it violated a phonotactic constraint of English, it was discarded. To this syllable, a prefix was randomly assigned. Again, if prefixation resulted in a combination which violated a phonotactic constraint, or if it was a real word, the word was eliminated. The corresponding derived term was formed by adding to the base term the suffix -*ity* and by altering the crucial vowel in the required manner. The resulting pair exemplifies a pairing of a base sound and a derived sound. All pairings of base vowels and derived vowels were made. Since there were five base sounds [ay, iy, ey, aw, ow] and five derived sounds [I, ɛ, æ, ʌ, ɔ], there were therefore 25 pairs of stimulus and response terms (the stimulus units). Only five of the six major alternations were included in the study. Since two of the back vowel alternations (uw/ɔ and ow/ɔ) have the same derived vowels, it was thought that including both alternations would result in a confusion as to the base vowel corresponding to the open o. The choice of ow/ɔ instead of uw/ɔ was arbitrary.

For each correct alternation, there were four incorrect ones -- incorrect in the sense that the derived sounds do not alternate with the base sound in English. These incorrect pairs served as controls. For example, for the correct exemplar subk[ay]pe-subk[I]pity, the following pairs exemplifying non-existing alternations served as controls: malp[ay]de-malp[ɛ]dity, purq[ay]te-purq[æ]tity, surf[ay]me-surf[ʌ]mity, exr[ay]ne-exr[ɔ]nity. The correct word pair and its controls constitute a group. In preparing the stimulus units for each group of word pairs, two restrictions were observed. First, no prefix was used in more than one alternation. Second, in the syllable that contained the crucial vowel sound, no consonant sound appeared more than once as the first consonant, and no consonant sound appeared more than once as the second consonant.

The words were submitted to three graduate students of linguistics who were native speakers of English and who knew the purpose of the experiment. They were asked to eliminate from the list forms which, in their judgment, were not possible English words, or forms which reminded them strongly of common English words, or which were too noticeable so as to facilitate learning. The list went through a number of revisions and revalidations before the final list was considered satisfactory (see Appendix A).

The words were then recorded on audio tape by a female native speaker of Canadian English. Using a PDP 12 computer, these recordings were digitized and re-recorded on LINC tapes. The actual stimulus materials and the test words were recordings from these master LINC tapes. This procedure permitted the experimenter to control for signal amplitude and to filter out excessive hissing sounds associated with certain sibilants. In recording the materials, proper randomization procedure of the stimulus units was observed.

Procedure

Ss were tested individually. The S was seated at a small desk. Four feet in front of him, on a high table, was an extension speaker, from which the materials were heard.

The S was told that the experiment was designed to find out more about general characteristics of human memory (see Appendix B for full text of the instructions). They were advised to treat the word pairs as adjective-noun pairs. In a familiarization pretest, S heard and repeated the stimulus words -- unpaired, i.e., all words of one type first (base or derived) followed by words of the second type.

Materials were presented aurally using the study-test method. In the study phase, all 25 stimulus pairs were presented. A three-second interval separated words of a pair, and a five second interval separated pairs. In the test phase, only the test words were presented, with a four-second interval between test words. The subject was asked to supply the word paired with the test word. The study phase and the next test phase were separated by a short high tone. Criterion was set at one error-free trial, or 10 trials, whichever was first. Subject's responses were recorded on tape.

Subjects were divided into two groups. One group heard the pairs with the base word ordered before the derived word. In the test phase, this group was given the base word as the test word, and they replied with the derived word. The other group heard the pairs in the order derived-base. The test word for this group was the derived word, and the subjects had to respond with the base word.

Design

Types of vowel sounds, i.e., whether base or derived, served as independent variables, with five levels for each variable. Each level of base vowel was paired with all levels of derived vowel, giving 25 treatment combinations, as shown in the following chart:

		Derived Vowels				
		I	ɛ	æ	ʌ	ɔ
Base Vowels	ay	ay/I	ay/ɛ	ay/æ	ay/ʌ	ay/ɔ
	iy	iy/I	iy/ɛ	iy/æ	iy/ʌ	iy/ɔ
	ey	ey/I	ey/ɛ	ey/æ	ey/ʌ	ey/ɔ
	aw	aw/I	aw/ɛ	aw/æ	aw/ʌ	aw/ɔ
	ow	ow/I	ow/ɛ	ow/æ	ow/ʌ	ow/ɔ

The alternations in the main diagonal are correct English alternations; all the other alternations are incorrect alternations and served as controls. The learning situation is essentially, then, in itself a transfer paradigm of this sort: A-B, A-{B, C, D, E, F}, where the first A-B is the pre-experimental linguistically determined linkage. The second A-B duplicates the first A-B; all other linkages are experimentally prescribed, and, furthermore, not in accord with linguistic usage.

A third independent variable is stimulus type, i.e., whether the subjects received the base form or the derived form as stimuli. Half of the subjects received the base forms as stimuli and the other half received the derived forms as stimuli. Assignment of subjects to stimulus type was random. All subjects did the task for all 25 vowel type treatment combinations.

The basic dependent measure was number of correct responses summed across trials. First correct response, last error trial, and error types were also analyzed. The expectation was that alternations mediated by some psychologically real generalizations should elicit more correct responses.

Experiment 2: Transfer

Materials

Twenty-five new non-existing but phonologically possible English adjective-noun pairs, of the same construction as the stimulus units in the paired-associate learning task, were used-as stimuli (see Appendix A).

Design and Procedure

The design and the independent variables were the same as those of the acquisition experiment. The general procedure was to present the stimulus terms only, and the subjects were asked to supply the corresponding associates. The stimuli were presented one at a time, with a four-second interval between stimuli, and in three blocks of presentation. The order of the stimulus words was varied in each presentation. (See Appendix B for full text of the instructions.)

The purpose of this test was to provide data on the transferability of the associations learned. Differential transferability was predicted between internally guided alternations and alternations linked by mediators developed only as a consequence of the experimental learning task. If in fact the generalizations that link real alternations are part of the speakers' mental grammar, then such generalizations should always be available for productive transfer to novel forms. The data will also be relevant in determining whether the subjects learned the paired-associates as individual paired units, making use of properties of the units that are irrelevant to the generalization being tested. In such a case, productive transfer to novel forms is an unlikely consequence.

Post-Test Interview

After the transfer test, the subjects were asked about their strategy in learning the word pairs. Nearly all of them could not articulate their strategies ("I just tried to remember the words"), and this was taken to mean that, the strategies (if any) were beyond the level of the conscious. The one or two subjects who said that they started out by trying to associate the stimulus words with real words, had to abandon this strategy because the words came so fast that they just had to concentrate on the stimulus words. Subjects were asked if they thought about word pairs like *divine-divinity*, *profane-profanity*, etc., or recognized an association between these real word pairs and the stimulus pairs; not one answer was in the affirmative. The ten introductory linguistic students were asked about whether it occurred to them that they were being tested on the vowel shift rule; none of them knew what the vowel shift rule was about.

CHAPTER FOUR

RESULTS

Data-analysis was guided by the following questions: Are correct English alternations easier to learn than non-English alternations? Do English alternations interfere in the learning of non-English alternations? Are there any other sources of learning interference? Is knowledge of correct English alternations easier to transfer to similar situations than knowledge of non-English alternations?

Tabulation of Raw Scores

Table 1 (tables and figures are in Appendices C and D, respectively) shows the responses summed over trials over subjects. The left half of the table contains front formation scores, and the right half tabulates Back Formation scores. In each formation half, the five squares of scores represent stimulus types. Rows represent required responses, and the columns represent actual vowel responses. Thus the first cell entry, 154, is the total number of times [I] was used as response to the stimulus [ay] when the required responses was [I], summed over all ten trials and for all the 20 front formation subjects. The maximum possible value of a cell entry is 200.

The diagonals in each square represent correct test responses. A correct test response was one in which a subject's actual response matched the required response. All off-diagonal values are, therefore, errors with respect to the requirements of the test. Of these error scores, some are correct responses with respect to English (the underlined scores), i.e., they are the responses predicted by the English vowel alternation generalization. The correct test scores represent learning scores and the correct English scores represent part of the overall interference scores.

Since the major concerns of this study are the extent of learning and the amount of interference with this learning induced by English grammatical habits, two summary tables that give indications of learning and interference are presented.

Table 2 is a summary of actual responses as a function of stimulus type. Entries in the main diagonal represent correct English scores. The off-diagonal values in each row are scores corresponding to the non-English or control alternations. It will be noticed that, except for the fourth correct English alternation, the values corresponding to the correct English alternations are larger than any of the values of the control conditions, and this is true in both front and back formation conditions.

Table 3 is a summary of actual responses as a function of required response types. Entries in the main diagonal are correct test scores. The greater magnitude of these scores over off-diagonal values suggests that, if frequency of responses is to be a measure of learning, subjects were in fact learning the alternations, both English and non-English. The overall total scores (bottom row of Table 3) represent the number of times the vowel types were selected as responses.

A Preliminary Overall Analysis

To provide a general view of the nature of the data, an overall analysis of variance was performed. The design of analysis used was a mixed model seven-factor ANOVA, with partially repeated measures, and with the last four factors embedded within the first factor. The factors are as follows: Formation Type (two levels: Front and Back Formation), Trial (ten levels), Stimulus Type (five levels), Required Response (five levels), Actual Response (five levels), and Subjects. To provide an estimate of within cell variation, the subjects were divided into four subgroups, with five subjects randomly assigned to each subgroup. Each subgroup was treated as an observation. Thus, the formation types had four observations each. The results of this analysis are summarized in Table 4.

As shown in Table 4, all interaction effects, except for three second order interactions, were significant at the .001 level. The three exceptions are Trial x Stimulus Type (n.s.), Trial x Required Response (significant at the .05 level), and Stimulus x Required Response (significant at the .01 level). This general result is as expected, since learning scores should improve through practice. The differential scores induced by specific types of stimulus and required response pairings are to be expected if it is to be demonstrated that correct English alternations have greater relevance to learning than non-English alternations. It now remains to be seen whether these correct English alternations constitute the class of alternations that induced the greatest amount of learning.

Post-hoc comparisons of groups of means were performed using Scheffe's technique. The comparisons made, the confidence intervals, and the significance probability levels are shown in Table 5. All comparisons, except two, were significant at the .01 level. The two nonsignificant comparisons corresponded to the comparison between the correct English back formation task ʌ/aw and its controls, and the overall front-back comparison between this alternation and its controls. In the Group A comparisons, the five correct English alternations were compared with the 20 incorrect alternations. Three comparisons were made: Front Formation, Back Formation, and combined overall Front and Back Formation. All three comparisons yielded significant values in favor of the Correct English category. This suggests better learning of the correct English alternations, irrespective of the formation task under which the alternations were learned.

In the Group B comparisons, each correct English alternation was compared with its control alternations, and, again separately for Front and Back Formation, as well as combined Front and Back Formation. All correct English alternations were significantly more learnable than their controls, with one exception: the fourth alternation, aw/ʌ, did not appear to be significantly more learnable than its controls when the task was Back Formation. This nonsignificant value resulted in a nonsignificant overall comparison value for this alternation.

The significant comparison values in Group C, in favor of Correct Test, suggest that the subjects did in fact learn the alternations, whether these alternations were Correct English or not. Because of these significant values, the significant value~~s in~~ the Correct English-Incorrect English comparisons (Group A comparisons) now appear to be the confounded effect of Correct English and Correct Test, since some correct English alternations also belonged to the Correct Test category. The comparisons in Group D were corrected for this confounding. The comparisons involve Correct Test categories that were at the same time Correct English, against Correct Test categories that were at the same time Incorrect Test; the effect of test is thus held constant. The comparison values are all significant, suggesting that 'pure' Correct English did in fact induce better learning than Incorrect English alternations. In the Group E comparisons, the effect of Correct English was held constant, thereby providing values that are purely Correct Test. The significant comparison values are in favor of Correct Test.

The comparisons in Group F involve the categories Peripheral vowels [I, ɔ] versus non-Peripheral vowels [ɛ, æ, ʌ]. Although Front, Back, and Overall comparisons were made, only the Front Formation comparison can be interpreted, since the dimension Peripheral is difficult to apply in complex vowels. The significant Front Formation comparison value suggests that alternations whose target vowels are at the extreme end of this dimension induced better learning. This point is discussed in some detail below.

Prior to data analysis, a decision was made that, to answer more specific questions, separate analyses of the front formation data and the back formation data would be necessary. This decision was made on the basis of the observation that, given the nature of the experimental task, the front formation task was not really comparable with the back formation task. In the front formation task, the stimuli were complex vowel sounds (e.g., [ay]), and the required responses were simple vowel sounds (e.g., [I]). In the back formation task, on the other hand, the stimuli were simple vowel sounds and the required responses were the complex vowel sounds. It is not unlikely that memorization of complex sounds is more difficult than memorization of simple sounds. The significant formation type main effect, in favor of front formation, lends support to the correctness of this conjecture. It should be noted that separate Back and Front analyses do not lead to a weakening of the interpretation of the results in the way that the main concerns of the study are jeopardized, but only that separate analyses require that certain conclusions may not generalize across formation types.

Analysis of the Learning Scores

The first set of analyses addressed the question of the learnability of the correct English alternations in comparison with the control, incorrect English alternations. Three learning indices were established: total correct test responses summed across trials, first correct response trials, and last error trials. Separate analyses were made for each learnability index.

The input data to the first analysis were the response frequency scores (correct test scores). The response frequency score of an alternation condition consisted of the number of successful matches made in ten attempts to learn the correct pairing. A three-factor ANOVA was performed, with Stimulus Type, Required Response Type, and Subjects as factors. The design used was a mixed model (fixed AB, random C), with repeated measures on the first two factors. The results of the Front and Back Formation analyses are summarized in Table 6.

In both Front and Back Formation analyses, the interaction between Stimulus Type and Required Response is significant (Front Formation $\underline{F}$ = 12.74; Back Formation $\underline{F}$ = 11.44, with $\underline{p} < .001$ and $\underline{df}$ = 16,475 for both ratios). This suggests that the learnability of the word pairs depended on the type of vowel sounds paired. Figure 1 shows that the alternations that resulted in the greatest learning in each group of alternations are the correct English alternations (a group of alternations consists of the correct English alternation and its four controls). The values plotted are cell deviation scores (a cell deviation score is the value cell mean minus the interaction means plus the overall mean). In both Front and Back analyses, and in all groups of alternations, the correct English alternations have higher deviation scores than their respective controls. This suggests that, as a class of alternations, the set of alternations that have real English examples induced more learning than the class of alternations that have no examples in English. Further support for this general finding is provided in the analyses of the first correct trial and last error trial scores.

The first correct trial score is the trial number in which the required response was first given. If none of the responses for all ten trials was correct, a value of 11 was assigned. The last error trial score is the trial number in which the last error occurred. If all responses were incorrect, a value of 0 was assigned; if all responses were correct, a value of 11 was assigned. A low score in both measures indicates early learning.

Separate analyses of variance were made on these two measures. The design of analysis was similar to the one employed in the previous analysis. Results are summarized in Tables 7 and 8. In all of these analyses, the interaction between Stimulus and Required Response is significant. The *F*-ratios and the probability values are as follows, with *df* = 16,475 in all of the comparisons: First Correct Trial, Front Formation *F* = 8.44, *p* < .001; First Correct Trial, Back Formation *F* = 2.0, *p* < .01; Last Error Trial, Front Formation, *F* = 8.61, *p* < .001; Last Error Trial, Back Formation, *F* = 3.52, *p* < .001.

These significant interactions suggest that early learning and persistence of learning errors are a function of the type of alternations being learned. Figures 2 and 3 plot the cell deviation scores of the alternations. In all but two exceptions, the deviation scores of the correct English alternations are lower than their respective control conditions in both First Correct and Last Error analyses, suggesting _early_ learning of the word pairs exemplifying correct English alternations. The two exceptions are the deviation scores of the [ʌ/aw] alternations of the Back Formation group in both First Correct and Last Error analyses. These values are exceptional in that the First Correct deviation score is higher than three of the four control conditions, and the Last Error deviation score is higher than all of the control conditions. With these two exceptions, the evidence provided by the First Correct and Last Error scores on the superior learnability of correct English alternations over incorrect ones, while still impressive, is not as overwhelmingly conclusive as the evidence provided by the overall learnability frequency scores.

Analysis of the Correct English Interference Scores

Error scores represent learning interference. Part of these interference scores consists of correct English responses. These correct English scores have been transformed into percentages of the total learning error scores for each row (see Table 9). For example, in the Front Formation task, when the stimulus was [ay] and the required response was [æ], 81% (the first entry in Table 9) of the errors were responses that were the correct English alternant of [ay], which is [I]. There are no entries along the main diagonal since, for these conditions, the correct English responses were also the correct test responses. An examination of these values reveals that in many cases, correct English interference scores constitute at least 50% of the errors, and only the scores corresponding to the [aw]-group of the Front Formation task and the score for the [ɛ/ow] Back Formation alternation went below the 30 per cent level. Since one of the main concerns of this study is the possible inhibitory effect of grammatical habits in the learning of novel linguistic patterns, direct comparison of correct English and correct test scores was made. The basis for this comparison, discussed in some detail in Chapter 2, is briefly recapitulated here.

If the subjects "know" the vowel alternation generalization in the sense that they employ it as a directing rule, the expectation is that subjects' memorization of the word pairs will be differentially affected. In particular, memorization of the word pairs that exemplify English alternations should be facilitated and memorization of the word pairs that do not exemplify real English alternations will be impeded. The strongest confirmation of such an expectation is, of course, a response pattern in which the non-English alternations are never learned and only the correct English alternations are learned, and furthermore, that all of the response errors consist of correct English alternant sounds. However, to expect to get this ideal result is to assume that there is a rule of language so strong that it cannot be temporarily suppressed in an experimental situation by intelligent, highly motivated university students. And clearly, this is an unreasonable assumption; students can be taught to memorize a reasonable list of arithmetical inequalities such as 2 + 2 = 6. Hence, a moderate expectation was made, which was that interference, if any, would be strongest in the first trials, and decrease as trials progress. Testing for this expectation requires comparison of subjects' correct English and correct test scores across trials.

A five-factor ANOVA was performed, with the following factors: Trial (ten levels), Stimulus Type (five levels), Required Response (five levels), Correct English/Correct Test, and Subjects. Subjects were grouped into four with five subjects randomly assigned to each group; each subgroup was treated as an observation. The results are summarized in Table 10.

In both Front and Back Formation analyses, the interaction between Trial, Stimulus Type, Required Response, and Correct English/Correct Test is significant (Front $\underline{F}$ = 1.28, Back $\underline{F}$ = 1.03, with $\underline{p} < .001$ and $\underline{df}$ = 144,1500 for both comparisons).

Figure 4 plots the cell deviation scores of the categories Correct English and Correct Tests for all trials. Because the values plotted represent deviation from cell means, and because there are only two categories, the line corresponding to one category is the mirror image of the other. The relevant information in this figure is the superiority of Correct English in the first four trials. Thus, in both Front and Back Formation, the values for the first four trials are positive for Correct English and negative for Correct Test. Starting with the fifth trial, when presumably learning started to take place, the values are reversed, i.e., Correct English scores are below the mean and Correct Test scores are above the mean. This is precisely the kind of results that would suggest the inhibitory effect of the internalized linguistic patterns being tested. Subjects were making errors in the first four trials in the direction of Correct English alternations; in other words, at a stage when learning has not taken hold yet, the nature of the response was largely determined by related linguistic experience.

Another Source of Interference: Phonetic Distance

The correct English scores do not account for all of the error scores; there are residual errors. It is possible that these residual error scores are not random but that they represent other sources of interference. An examination of the data suggests that two very likely sources of interference are phonetic similarity and phonetic polarity ("oppositeness").

Table 11 summarizes these non-Correct English error scores as a function of Stimulus Type (rows) and Actual Response Type (columns). Table 12 is a summary according to

Required Response (rows) and Actual Response (columns). For reasons that will be made clear later, discussion will be confined to the front formation data. The total error scores (bottom row of Table 12) represent the number of times the actual response vowel types were erroneously given to be the required (correct test) vowel response.

It will be noticed that the vowel types that provided the greatest interference were [I, ɔ]. If the totals were to be plotted, a *u*-shaped line would fit the points. This suggests that subjects were making errors in the direction of the endpoints along a front-back vowel dimension. These "endpoints" will be referred to as "peripheral" sounds; the other vowel sounds are "non-peripheral". (The term "peripheral" is used as a matter of expository convenience; it is not suggested here that the basis of the distinction in the minds of the subjects is articulatory in nature.)

If the influence of Stimulus Type is taken into consideration (Table 11), it will be seen that this tendency towards peripheral sounds varies according to stimulus type. Figure 5 plots the Front Formation values of Table 11 (the zero values along the main diagonal are "skipped over"; furthermore, certain sharp corners of the lines have been smoothed out). The stimulus sound [ay] resulted in a preponderance of errors towards the target configuration [I]; the stimulus [iy] produced a greater incidence of non-English errors in [ɔ]; [ey] is nearly level, and both

[aw] and [ow] gravitate toward the front end.

The striking aspect of this set of data is the pronounced tendency of the error trends of the sounds [iy, aw, ow] towards phonetic polarity. That is, the error was in the direction opposite to the target sound: back for the front targets and front for the back targets.

If, now, the effect of Required Response (Table 12) is taken into consideration, a slightly different picture emerges. The response pattern for the "middle" vowel [æ] is nearly level; it thus duplicated the response pattern of its alternant. But now the error patterns of the vowels [ɛ] and [ʌ], whose counterparts have shown a tendency towards phonetic polarity, show a tendency towards phonetic similarity. That is, the preponderance of errors is towards high front for the front [ɛ], and towards low back for the back [ʌ]. These two vowels are closest to the peripheral [I] and [ɔ], respectively, and it appears that errors were pulled in the direction of the peripheral sounds, or to the sound more similar with respect to the opposition front/back.

What this suggests is that, independent of the influence of correct English patterns, responses were further influenced by sound similarity and sound distinctiveness. If this is a general trend, it is to be expected that such patterns of errors should also be evident in the case of correct test responses. That is, analysis of correct test responses should reveal that subjects view the peripheral vowels to be somehow distinct from non-peripheral vowels, and, furthermore, that discrimination of the actual response types is also being made on the basis of frontness/backness.

A discriminant function analysis using the direct method solution (see Nie, et al., 1975) was performed on the correct test scores. Subjects served as variables and the groups to be discriminated upon were five alternation groups, the alternations being grouped according to required response type. Discussion will be confined to the front formation analysis (however, a similar analysis of the back formation data was made and the results are here reproduced). The relevant results are in Table 13.

Four discriminant functions have been identified. Only the first two, however, have significant chi-square values. The group centroid loadings for the first discriminant function distinguish between the groups [I] and [ɔ] on one hand and the groups [ɛ], [æ], and [ʌ] on the other hand. The distinction is between peripheral and non-peripheral vowels. The second discriminant function distinguishes between [I, ɛ, æ] on the one hand, and [ʌ, ɔ] on the other. This dichotomy corresponds, of course, to the distinction front-back. The results thus lend support to the idea that subjects' response patterns were indeed influenced by properties independent of, and in addition to, properties of the English vowel alternations being tested.

The decision to restrict interpretation to front formation data in this section will now be explained. It will be noted that the results of the front formation analysis have been presented in terms of a somewhat restricted vowel space. In this vowel space, the front formation response vowels may be distributed along a front-back dimension, thus: [I - ɛ - æ - ʌ - ɔ]. The back formation response vowel sounds, being complex, do not lend themselves to such a straightforward distribution. Any such attempts should take into consideration two aspects of the vowels--properties of the initial sound ([a] in [ay]) and properties of the terminal sound ([I] in [ay]). The problem here is that it is not known which aspect of these complex sounds played a salient role in learning, hence, the basis for a meaningful distribution of the sounds is difficult to establish.

Indeed, the original motivation for including the back formation condition was the somewhat overly optimistic expectation that by contrasting front formation and back formation, something of interest could be said about the reality of the direction of derivation. The realization that the two conditions may not be comparable in a learning-memorization experiment came only after the data had been gathered. It must be noted, however, that non-comparability of the alternates is only a possibility, and is by no means a certainty. However, as long as this possibility exists, any comparisons made are uninterpretable.

Analysis of the Transfer Scores

The transfer part of the experiment was predicated on the assumption that subjects would learn all the alternations in the allotted time, or, in other words, that subjects would have knowledge of the alternations in "equal amounts" at the start of the transfer task. This assumption was necessary in order for the transfer scores to be treated as transfer scores. A transfer test measures transfer of learning, and if learning took place at quantities that varied from alternation to alternation, transfer would then be a function of whether or not an item has in fact been learned, and not whether learning is strong enough to form the basis of a production strategy for new materials.

The assumption of equal knowledge of the alternations was not supported by the eventual results of the learning experiment (only one of 40 subjects learned to criterion). The results of the learning test show that the alternations were learned in differential quantities, generally in favor of the correct English alternations. Thus, at the start of the transfer test, the subjects knew some alternations better than other alternations. This point precludes analysis of the transfer scores as transfer scores.

Although the transfer scores cannot be treated as measures of transfer of learning, something is to be gained in an analysis of the scores. The transfer items were new nonsense words, and the transfer test procedure eliminated the study (feedback) phase of the learning test. The test, therefore, may be considered as a modified production test--modified in the sense that the subjects, on the basis of experience in the learning phase, were aware of the response possibilities (as opposed to the standard production test in which response range is relatively unrestricted). Making a response, then, amounted to "selecting the best answer".

Analysis of variance was performed on the data, separately for Back and Front Formation. As Table 14 shows, in both Front and Back analyses, the only significant main effect is Actual Response ($\underline{p} < .001$, $\underline{df} = 4,2375$), and any of the interaction effects involving this factor were significant at the .001 level. A look at Figure 6, which plots the cell deviation scores of the Stimulus Type x Actual Response interaction, shows that subjects made more correct English responses in all groups of alternations except in the fourth group (the aw/ʌ group) of the Front Formation task. Thus, while the transfer scores do not provide new information, they provide a confirmation of the findings of the learning experiment.

CHAPTER FIVE

DISCUSSION

The general findings of the experiment have been that nonsense word pairs exemplifying correct English alternations induced more learning than nonsense word pairs that do not exemplify correct English alternations, and, furthermore, that the relative low learnability of the non-English alternations may be attributed to interference due to knowledge of certain aspects of the morphology of English, as well as interference due to certain phonetic properties of the sounds to be learned. These findings will be discussed as they relate to certain topics in verbal learning and interference, the organization of the lexicon, and the issue of the psychological reality of linguistic constructs.

On Two Sources of Interference

Subjects' attempts to learn the experimentally prescribed alternations that do not duplicate English alternations were impeded by two forces. The first is the pressure of pre-experimentally learned, language-prescribed alternations, which competed against the 20 non-English alternations. This type of interference will be referred to as *morphological interference*, since the source is morphological knowledge. The English alternations are morphologically determined, not phonetically conditioned. That is, the pairing [ay]-[I], for example, as in *divine*-*divinity*, is arbitrary; there is no phonetic necessity for such a pairing. The second source of interference refers to certain phonetic properties of the target sounds, thus, the term *phonetic interference* will be used. It will be suggested that phonetic interference is a function of vowel systems in general, whereas morphological interference is language-specific.

Morphological interference. One way to account for the facilitated learning of the correct English alternations is to start with the assumption that, as a result of experience with similar materials, subjects had pre-experimental 'knowledge' of the English alternations. The words "similar materials" should, of course, be interpreted to refer not to the nonsense words themselves but to the crucial vowel sounds. What this assumption leads to is a bi-partitioning of the stimulus materials into two types: alternations that subjects already 'knew' and alternations that they had to learn for the first time.

If the materials are so partitioned, it would follow that the subjects performed not one but two tasks. The first task was the recall of associations already known, a recall task that was made simple by the cueing function of the test word. The other task was the acquisition of new associations, a task that in itself was more difficult than simple recall. Acquisition of the new associations was rendered even more difficult in that, in contrast to the usual experimental transfer situation, the old associations were not to be extinguished, but were in fact to be retained and used as bases for generating responses. Competition between 'old' and 'new' responses was, therefore, active throughout the experiment. The degree of competition induced by this pre-experimentally learned set of responses was extensive, which resulted in massive interference in the learning of new associations. In fact, only one of 40 subjects learned to criterion. In other words, linguistic knowledge proactively inhibited the learning of new linguistic patterns.

As a consequence of the above two-task restructuring of the experimental situation, the experimental paradigm became the transfer schema A-B, {A-B, A-C}. The first A-B is the learning phase, which is pre-experimental. The second set of linkages is the test phase, where (the second) A-B represents the correct English alternations and A-C the non-English alternations. Between A-B and A-C, the former should be easier to manage, since it is an exact copy of the pre-experimental condition. It can be said that while the subjects had all of their lives to learn A-B, they had no more than an hour to learn A-C.

Since morphological patterns of the type similar to English vowel alternations are language specific, morphological interference must perforce be language specific. In other words, only speakers of a language which exhibits the morphological basis of the interference are susceptible to this type of interference. If the present test were to be given to, say, French speakers who have not had extensive exposure to English, interference from this source should be nonexistent. Phonetic interference, however, should, under this view, still operate, since, as will be pointed out in the next section, this type of interference appears to derive from property of vowel sounds in general.

Phonetic interference. An unexpected but interesting finding from the analysis of the front formation data concerns the tendency for the peripheral sounds [I, ɔ] to be the preferred responses. They were easier to learn and subjects made more errors in their direction. In terms of overall frequency, the values are as follows: [ɔ] = 1232, [I] = 1017, [ɛ] =949, [æ] = 900, [ʌ] = 675. Why should an ordered relationship of this sort obtain among the vowels?

One possible explanation is frequency of occurrence of these sounds in the language. The most frequently occurring sounds may achieve a degree of prominence that manifests itself in a kind of mental set or in the form of a readily-activatable set of well-developed motor habits, so that these sounds are much more readily available as response items than infrequently occurring sounds. If this is true, it is to be expected that the peripheral sounds occupy the highest ranks in a frequency count of English vowel phonemes.

One such count is Dewey's (1923), reproduced in Miller (1951, p. 86), which was made on a phonemic transcription of standard English prose. The ranking of the five target sounds included in this study are as follows (in per cent of total sounds in the corpus): [I] = 8.53, [ɔ] = 4.56, [ɛ] = 3.95, [æ] = 3.44, [ʌ] = 1.33. (Note: the value for [ɔ] is the total of the values for three low back variants, [ɔ], [a], and the inverted 'a'. The motivation for collapsing these vowel categories is the fact that the subjects of the present experiment spoke a variety of English in which all these categories are in fact collapsed into the one category [ɔ]). Voelker's (1935) count, also reproduced in Miller (p. 145) as part of a table by Irwin (1948), duplicated Dewey's results (in per cent of the total number of vowel sounds in the corpus): [I] = 20.56, [ɔ] = 16.43, [ɛ] = 8.06, [æ] = 7.98, [ʌ] = 1.12 (again, the value for the low back vowel is a composite of the values corresponding to the "inverted a" and the "open o"). These frequency rankings and the experimental ranking do not match in all of the details, but they agree in general outlines. The peripheral sounds are on top, the stressed schwa is lowest, and the nonhigh front vowels are intermediate. Neither the Dewey nor the Voelker count, however, is ideal because the reduced vowel (the schwa) was treated as a separate category; the schwa in fact received the second highest ranking in both counts (4.63% in Dewey's, 17.76% in Voelker's). Further support is provided in a more recent count (Denes, 1963), where the following values are given (in per cent occurrence of total number of phonemes): [i] = 8.25, [ɛ] = 2.81, [æ] = 1.52, [ʌ] = 1.67, [ɔ] = 3.50 (the value corresponding to the last segment is again a composite of three low back vowels, the key words to which are "bought", "pot", and "part".)

The primacy of the peripheral vowel sounds may be a property of vowel systems in general. In frequency counts of vowel phonemes of languages other than English, the peripheral sounds come out nearly always on top. Greenberg (1966, pp. 18-19) reproduced some of these statistics. In five of seven languages sampled, the low back sound is consistently highest, and always followed by the high front vowel. In the two exceptions, Czech and Hungarian, [e] occurred most frequently, followed by the low back and the high front vowels. In a count involving 29 languages, Kramsky (1966) reported the following frequencies: [a] = 13,356, [i] = 7,359, [e] = 5,673, [u] = 4,507, and [o] = 4,384. And, again, the peripheral sounds have the highest frequencies.

Indeed, support for the primacy of the peripheral vowels comes from other fields. In the area of child language, Jakobson (1968) observed that peripheral vowels are learned first. Jakobson predicted that the sounds most likely to be lost or disturbed in aphasia are those sounds that are acquired last, in other words, in the present case, non-peripheral sounds. It is noteworthy that no report has as yet been made in the literature on speech dissolution on the loss of the peripheral sounds ahead of non-peripheral sounds. In speech perception, the peripheral (corner) sounds provide maximal contrast and are "acoustically stable", two requirements for effective perceptual distinguishability (Liljencrants & Lindblom, 1972).

The theory of markedness (see Chomsky & Halle, 1968, chap. 8) appears to provide a linguistic basis for the priority given to peripheral sounds. In markedness theory, the least marked sounds are the least complex in formal terms. It happens that, of the target sounds, the low back and the high front vowels are the most unmarked, hence, the least complex. If primacy is defined as a function of least markedness or least complexity, then the greater response availability of the peripheral vowel sounds finds a formal representation in markedness theory.

In Chomsky and Halle's system, the ranking of the five target sounds are as follows (numbers correspond to the complexity indeces; see Chomsky & Halle, 1968, p. 409): [a] ([ɔ] in our system) = 0, [I] = 1, [ɛ] = 2, [æ] = 2, [ʌ] = 3. In Chomsky and Halle's system, the nonhigh front vowels are of equal complexity; in the experimentally determined ranking, the difference between these two sounds is not significant. The fit between the formal based and the experimentally based complexity rankings is perfect.

Using many more phonetic parameters, Schane (1973) came up with a ranking of the target sounds essentially similar to Chomsky and Halle's, except that, with respect to the sounds of interest here, in Schane's system, [ɛ] came out less complex than [æ]. Thus, if absolute frequency values are to be the basis of ranking, Schane's system may be said to provide a realistic fit of the experimental data.

Markedness appears to be a variable not only in learning, but also in at least one other type of mental activity. An experimental demonstration of the perceptual reality of the markedness of certain vowels has been reported by Terbeek (1973). On the basis of similarity judgments of triplets of vowel sounds, Terbeek concluded that "mid, low front, and central vowels are all perceptually different as a class from 'corner vowels' [i.e. high front, low back, and high back vowels]" (p. 676). The first two of Terbeek's corner vowels are our peripheral vowels; the third corner vowel was not included in the present experiment. The corner vowels are all marked, in Terbeek's terms, for tongue position. Terbeek's subjects were native speakers of five different languages, suggesting the 'universality' of the effect on perception of peripheral vowels.

In the next section, the findings will be discussed in relation to two general models of verbal learning and memory. The point of interest is the possible relevance of the findings in clarifying certain aspects of the models;

or, conversely, the usefulness of the models in providing an explanation for the findings.

Relevance of the Findings to Two Models of Verbal Learning

Stimulus-response models. Stimulus-response (S-R) associationistic models view learning as the acquisition of connections, or associations, between a given stimulus and a particular response. The stronger the bond between stimulus and response, the greater the probability that the stimulus will elicit response. In classical S-R models, learning an S-R association is direct, that is, no event mediates in learning. In later models, e.g., Hullean models, mediating events are postulated to intervene between S and R. Learning is promoted if the following conditions are present: the S-R units occur contiguously, repeatedly, and the elicitation of the response unit is followed by a reinforcing event. (See Kausler, 1974, chap. 1, for a review of these models)

The greater frequency of correct English responses in the present experiment can be accounted for in terms of the S-R notion of "associative strength". Since the connections between correct English S-R units have been established pre-experimentally, whereas the connections between non-English S-R units were being established only in the course of the experiment, the association between English S-R units were stronger than those of non-English S-R units. In responding to the stimuli with correct English alternations, subjects were not acquiring a new response category; they were merely trying to retrieve this category. As pointed out earlier, the distinction is between recall of previously learned responses and the learning of new ones.

But notice that what is being explained here is the activation or recall of already existing associations. When the question of the pre-experimental acquisition of these correct English associations is brought up, classical S-R models can provide no adequate explanation. Classical S-R models have no mechanism to explain how the associations have been learned in the first place. The reason is that a basic requirement for learning a connection is that the S-R units occur contiguously and repeatedly. However, it is unlikely that the S-R units (i.e. their correct English analogues) have occurred contiguously and repeatedly in the past linguistic life of the subjects, and that these units have been recognized unequivocally as stimulus-response units. It is unlikely that the unit [ay], for example, has co-occurred with the correct alternate [I]; the word divine is not always uttered or heard as a stimulus to, or a response to, or in the context of, the word divinity. The association between the correct English S-R units cannot, therefore, be direct. This is a point that Braine (1974) apparently missed, when, hypothesizing on how alternations of this sort are naturally learned, he wrote that "the learner's pattern-learning mechanism is such that he registers alternations between phonemes that occur in particular environments when the alternation pattern occurs frequently enough" (p. 293). It is not to be denied that the alternation patterns occur in the language, but the point is that the alternates seldom occur contiguously and often enough for a direct association between them to be registered. Indeed, the non-occurrence of a direct associationistic linkage between stimulus and response units is a major difference between language learning and behavior in the experimental laboratory and in a natural setting, a point which Chomsky (1959) raised in his review of B. F. Skinner's Verbal Behavior. (Chomsky in fact raised the much more serious problem of the difficulty in identifying the appropriate stimulus units for particular linguistic responses.)

A departure from classical S-R models is the postulation of events that intervene between stimulus and response. In Hullean learning theory, this mediating event takes the form of a response-produced stimulus. Thus, the chain of association is: S - [r - s] - R, where the brackets signify the "internalized" segment of the chain (Kausler, 1974, p. 9). Let us simplify the internalized segment of the chain into a small r, which itself serves as the stimulus to R. In this simplified version, the associative chain for the pair KING - TABLE may be: KING - Arthur - TABLE. Thus, this mediational theory makes it unnecessary for KING and TABLE to occur contiguously; it is only necessary that both occur contiguously with a common item, in this case, ARTHUR. This concept of mediated learning may be used as a basis for developing an argument for the explanation of associations between natural language stimulus and response units that do not occur contiguously with any degree of frequency.

Suppose it is assumed that, in linguistic usage, when a word is encountered, all its derivatives are by indirect mediated association also encountered. Encountering _divine_, for example, involves a 'coming to mind' -- certainly not consciously -- of the derivatives of _divine_. The chain of association may be: Divine - r - {DIVINITY, DIVINATION, DIVINER, etc.}. That is to say, _divine_ elicits, via a certain mediator, multiple responses. Multiple response potential is well-motivated in S-R psychology, in the concept of "habit-family" (cf. Kimble, 1961), which is the set of responses to a given stimulus. For example, the stimulus TABLE may have a habit family consisting of CHAIR, FOOD, DESK, OFFICE, etc. Clearly, this habit family is semantically based. It is a possibility that the derivatives of a word constitute a habit family, that is, a syntactic (derivational) habit family. Thus, a reasonable S-R based solution to the problem of non-contiguous occurrence of stimulus-response units may be developed. Unfortunately, it creates a new problem: What is the nature of this small _r_?

Is _r_ the semantic core of the set of derivationally related words, the phonological core, or a combination of both, that is, the lexical entry (in the generative grammatical sense) /divin/? This problem does not arise in the chain KING - Arthur - TABLE, since the mediating response has an identity distinct from the stimulus and the response. In a chain such as DIVINE - r - DIVINITY, it seems unavoidable to postulate an entity abstracted from properties of the stimulus and the response. But once this is done, the concept of input transformation enters into the picture. The stimulus *divine* is transformed into whatever is the form of *r*, and *r* is re-transformed into the response *divinity*. It is not clear to what extent such a process of input transformation, certainly a cognitive process, if allowed to enter into S-R learning theories, will destroy the spirit and identity of S-R theories as S-R theories. Bever, Fodor, and Garrett (1968), have argued that attempts to expand stimulus-response theories to accommodate cognitive control processes such as response selection and response suppression violate the meta-postulates that define S-R theories. In any case, input transformation is an essential property of information processing models of learning, which will now be discussed.

Information processing models of verbal learning. The following is a sketch of the salient properties of information processing (I-P) models of verbal learning and memory. For more discussions, see Norman (1969), and Kausler (1974).

Since learning occurs in finite time and the locus of learning (the brain) occupies finite space, the essential problem for information processing models of verbal learning is to account for the seeming infinity of the learning products. Information-processing models thus view verbal learning as a storage-retrieval problem. The view is taken that input information is processed (organized, integrated, etc.) and transformed into modalities that are efficient for storage and its subsequent retrieval. Information transformation is a very complex task. It is a multi-step process, and each step requires that certain tests and decisions be made. This decision-making aspect of the task is believed to be carried out by a central executive system which directs and controls processing.

A predominant type of I-P models is one that makes use of organization factors in encoding and in retrieval. Information is organized such that retrieval is by blocks of information and not in terms of individual items. That information in memory is so organized was argued for convincingly by Miller (1956) in his now classic paper on the limits of human memory. Miller set the limits of primary memory to 7 plus or minus 2 units. If this is all that primary memory can handle, how is it able to remember a sizeable amount of information? Miller introduced the concept of "chunking". Thus, a unit can be a chunk of information, organized in some fashion. Recovery of an item leads to recovery of other items.

One of the most powerful organization models of verbal learning is a generative model, which has the productive capabilities of an axiomatic system. Like an axiomatic system, a generative system postulates rules of formation that apply on given elements to produce well-formed strings (sentences). Such a model of verbal behavior has been proposed by Mandler (1967). Mandler's proposal may be viewed as the culmination of intersecting views from three fields. In verbal learning theory, there was the realization that learning a language cannot simply be the development of associative bonds between directly specifiable discrete stimulus and response units (cf. Miller, Galanter, and Pribram, 1960). At about this time, computer ideas were beginning to be imported seriously into psychology. In computer terminology, a productive system consists of input devices and output devices, mediated by a program which contains specific instructions on how the input data is to be transformed into the required output materials. And in linguistics, Chomsky (1957) suggested that a more adequate theory of language is a generative theory. Mandler wrote:

> Thus, for example, a simplified version of an adequate theory to account for the production of English consists of a set of rules that we usually call grammar or syntax, and a set of building blocks that we can call words.... The syntactic rules are applied to words in storage to generate sentences and intelligible speech. Problems of psycholinguistics are not of concern here; the major point is that generative theories may be generally useful to a psychology of thinking. (p. 16)

Mandler explained that the reasonableness of generative rules as a model for human thinking stems from the fact that given such rules, "not all associations need to be stored specifically" (p. 17). To Mandler, the usefulness of generative rules lies in their employment as "accessibility cues", and to the fact that "they make it possible to generate novel material instead of simply reproducing some previous input" (p. 28).

If now, it is assumed that speakers employ generative rules, the greater frequency of responses exemplifying correct English alternations observed in the present experiment is easily explained. Subjects had rules to use to generate these responses, but no rules to generate non-English alternation responses. But, again, as in associanistic theories, what is being accounted for is simply the activation (this time, generation) of known responses. And S-R theories do not fail in this regard. The problem for a rule-governed model such as Mandler's, and, in fact, for S-R theories as well, concerns the origin and development of the mechanism in terms of which behavior is explained. The associationist believes that a bond develops as a result of repeated occurrences of S-R contingencies; the inadequacy of this suggestion in accounting for natural language skills has already been pointed out. On the origin of rules, Mandler wrote:

> The origin of rules is a complicated question that has hardly been touched. One possible source is the development of so-called analogic structures, cognitive representations that develop as a function of the frequent occurrence of overt behavior or environment-behavior contingencies. These analogic structures permit cognitive, overt manipulation and may, in some instances, be prototype rules. (p. 19)

Mandler's account is essentially associationistic, with its reliance on the frequent occurrence of environment-behavior (i.e. S-R) contingencies. Hence, it, too, must explain how non-contiguously occurring S-R units are learned.

Apart from the problem of the origin of rules, organization theories have one other fundamental problem: They rely on the existence of an omniscient central regulatory mechanism, but no explanation is offered for the origin and development of this mechanism. This mechanism analyzes the input string, locates and identifies the locus of change, scans the list of rules, flags down the relevant rules, does the match between the input structure and the structural requirements of the rules being tested, and performs the indicated change. These are all very complicated processes, and the mechanism that carries them out is being assumed. This point is the basis of Postman's (1972) criticism of rule-governed or I-P models of behavior. He wrote: "These specialized components [the central processors] are preprogrammed to carry...out [the operations of learning and memory] and thus the most difficult theoretical questions are begged" (p. 18). The problem appears to be another instance of "infinite regress" (for some comments, cf. Neisser, 1967).

In summary, both associationistic and organizational models of learning can account for the greater frequency of responses exemplifying correct English alternations. Associationistic theories, which rely on contiguous, repeated, and reinforced occurrences of the S-R units cannot account for the initial learning of the alternations. Organizational theories postulate rules and control mechanisms that likewise are able to explain the findings, but no account is offered as to how these rules and control mechanisms become part of the knowledge of the individual. In short, S-R models are simpler, but inadequate; whereas organization models are no more adequate, but, at the same time, rather mysterious.

On the Organization of the Lexicon

One important issue in linguistic theory concerns the nature of the lexicon. Two aspects of this question that have received some attention are: First, how the units of individual lexical items are to be represented, and second, how the lexical items are organized (if at all). The findings of the present study do not decide on the issue; however, they provide a useful basis for clarifying certain positions.

In connection with the first point, Wickelgren's (1969a, b) model is the most clearly articulated. Wickelgren suggested that the units of linguistic storage are *unordered* context-sensitive allophones, which are joined together in a vast network of associative chains. Thus, the units of the item /dIvayn/ are stored as "#*d*I" (i.e., the allophone *d* in the context #_I), "d*I*v", "I*v*a", "v*a*y", "a*y*n", "y*n*#". These units are not ordered linearly as such in the lexical space. They are, however, connected by an associative chain, and their appearance in speech in the proper order is effected when the chain is activated.

The findings of the present experiment suggest that the context in terms of which an allophone is supposed to be prepresented cannot be as simple as Wickelgren suggested it to be. Let us recall that the subjects in the present experiment had to learn alternations whereby the sequence "-ayc#", for example, was made to alternate with each of the following sequences: "-ICity#", "-ɛCity#", "-æCity#", "-ʌCity#", and "-ɔCity#" (where C stands for a consonant). If the context of an allophone is defined only within a form, as Wickelgren suggested, the 'allophones [I, ɛ, æ, ʌ, ɔ] (allophones only in the restricted context of the requirements of the test) should have equal learnability. This is because these allophones all occurred in contexts conformable with each other, and the forms within which they were embedded occurred with equal frequency in the experiment -- conditions that should result in associative chains of equal strength. The findings, however, showed a marked preference for a particular element, in this case, [I].

It cannot, of course, be said that the chain for "-ICity" is stronger on account of the fact that it occurs in English (e.g., divinity), since all the other sequences also occur in english (serenity, profanity, etc.). If the idea of associative chaining of the units of lexical items in linguistic storage is to be retained, it must be said that a chain has been established between "-ICity" and some other element outside of the lexical form, which element is in fact the alternant "-ayC#". In other words, there is a pre-experimental linkage between "-ayC#" and -ICity", whereas none of the control forms have this linkage. Clearly, the sort of associative chaining required to insure that the correct allophones of a word form are selected and ordered in just the right way must extend to elements outside of the word form.

In another, less serious respect, Wickelgren's associative chaining model is inadequate. It does not express the fact that chaining can be established across a variable. Obnviously, the link is between the crucial vowel and the suffix "-ity", and not between the crucial vowel and the next consonant. The unit(s) between the crucial vowel and the suffix "-ity" is irrelevant in determining the proper shape of the vowel. (For further discussion and criticism of Wickelgren's model, see Baker, 1974).

On the organization of the lexicon, two related sub-issues may be identified: first, what items appear in the lexicon, and, second, on what basis are these items organized (syntactic, semantic, phonological, or some combination of these factors). In connection with the first sub-issue, two opposing views that will serve as anchor points will be defined: the fully generative approach and the morpheme list approach. In the morpheme list approach, the lexicon is a mental approximation of a desk dictionary; a lexical entry duplicates its corresponding dictionary entry. Full words as well as roots and affixes are entered in this mental dictionary. There are no rules that relate classes of words with other classes of words. The entries are completely unordered and unorganized. In an organization scale, a morpheme list lexicon is down at the bottom end.

At the other end of the spectrum is a fully generative lexicon. The lexicon consists of basic elements and rules of formation that generate other, non-basic, elements. Thus, divine is listed but not divinity; the latter is to be derived from the former through a de-adjectival noun formation rule. In this lexicon, each lexical entry is completely non-predictable from other entries in the lexicon, and rules are as maximally general as possible. The semantics and phonology of derived words are determined by the semantics and phonology of the basic word plus whatever information is added as a result of the application of rules. Nothing has been said about whether the minimal set of lexical entries has any organization; the high degree of organization refers to the set of the total words in the language. Such was Lee's position in 1960. It is assumed that the generative semanticists adopt a version, in fact, perhaps a stronger version, of this approach, since they propose to decompose certain lexical items into more primitive elements. Lightner (1975) espouses what amounts to this original position ("the lexicon contains only roots as entries" (p. 632)).

Between these two opposing views are a number of intermediate positions. Chomsky (1970) paved the way for the emergence of non-fully generative approaches. Chomsky suggested that NPs like *Bill's decision to go* should not be derived from S-dominated NPs like *Bill decides to go* but should instead be generated by base rules. In this view, *decide* and *decision* are treated to constitute a single lexical entry, unmarked for the syntactic features that distinguish verbs from nouns. Chomsky's position is still within the fully generative approach. However, Jackendoff (1975) adopted Chomsky's analysis of the above nominals, but disagreed with him on the representation of the items *decide/decision*, and proposed an alternative approach which treats these words as distinct but related lexical entries. In Jackendoff's view, these words are to be related not by generative rules but by lexical redundancy rules. Jackendoff's position is a partial-list approach in that certain 'derived' words, i.e., derived nominals, are to be listed, along with their related 'basic' forms.

Kiparsky (1971) hinted that a partial-list approach is a viable one; he pointed out, on the basis of Hale's (1970) Maori data, that Maori speakers appear to memorize certain surface forms that in a formal analysis are derived from more basic forms. Zimmer (1969) likewise suggested the possibility of a list lexicon. Hsieh (1970) went a step further and proposed that morpheme alternants are to be listed in the lexicon, along with information as to the exact occurrence in surface of the variants of a form, until the time it is discovered that the forms underlie a regularity, in which case, a rule is formulated and internalized, and the derived forms are presumably eliminated from the list. Hudson (1974) and Braine (1971, n. 12; see, also, Braine, 1974) came up with essentially the same proposal as Hsieh's; in Braine's system, "the learner would learn two grammars, in one of which the derived forms are generated by rule, and in the other by enumeration" (p. 181, n. 12). Finally, a fully word-list approach has been proposed by Venneman (1974), which will be discussed shortly. The trend evident in these approaches is to de-generativize the lexicon. Non-fully generative approaches differ as to the class of items to attribute to rules on one hand and to memorization on the other hand.

In all of these schemes, what is the function of generative rules? In the strong generative position, rules generate, as well as relate, forms. In non-fully generative systems, for example, Jackendoff's system, generative rules cannot be said to generate forms that are already in the lexicon. But forms like decide and decision must somehow be related in the lexicon, since this information is presumably part of speaker knowledge. Jackendoff proposed lexical redundancy rules that not only relate these forms, but also allow for the entry in one place of phonological and semantic properties common to both forms. (Jackendoff, of course, has to postulate generative rules as well, to handle alternants not listed in the lexicon.) In Hsieh's approach, rules likewise generate and relate forms, but only in so far as the speaker has discovered and internalized these rules. It is not made clear how the variants of non-productive alternations, which speakers feel are related, are to be related in Hsieh's lexicon. Hudson would write rules for these non-productive alternations, rules of the form "variant A occurs in environment X, and variant B occurs elsewhere".

Venneman's view of the function of rules is unique. He assigns them a dual function: a generative function and an organizational function. To Venneman, an individual's lexicon is a collection of all words (not morphemes) he knows, in addition to generative rules. Thus, divine and all its derivatives are listed separately in the lexicon, and rules do not generate these words, nor any other words that are already part of the individual's lexicon. Venneman's position is thus a fully word-list approach, and is only one step short of the morpheme-list approach in terms of the number of entries in the lexicon. The difference, however, is that Venneman's word-list lexicon is a highly organized lexicon.

If rules do not generate forms, what function do they perform? Venneman took away the generative function of rules with respect to existing words, but not the relating function. In his view, rules serve as organizing principles of the lexicon. This organization function of rules can only be in the psychological sense, since, elaborating, he says, "rules facilitate retention of the lexical items as well as of specific semantic, syntactic, and phonological properties of lexical items in the lexicon" (p. 370). Venneman, however, retained the generative function of rules with respect to incoming forms. Thus, he wrote that the generative use of rule applies "to the spontaneous creation of new words and their analysis, [and] the adaptation of foreign words" (p. 367). In other words, rules "permit the spontaneous production and the analysis of words not known (not in the lexicon), and facilitate their incorporation and retention in the lexicon" (p. 370).

The general finding of the present study that is relevant to the issue of lexical organization concerns the basis of speaker knowledge of lexical relatedness. The demonstration that this knowledge extends to meaningless but English-sounding words suggests that phonological information serves as a basis for lexical organization. The role of non-phonological information in defining this knowledge in this particular context is minimal. The subjects were provided with only the very general semantic information that the forms shared the same general meaning, and the syntactic information that they were adjective-noun pairs.

Since the above finding is obviously limited, only classes of hypotheses about the lexicon can be eliminated. First, the class of pure-list hypothesis, that is, those that do not admit of any organization of the elements, e.g., a morpheme-list approach. The results this approach would require for support are those that show that the materials have equal learnability. Of organizational hypotheses, those that rely on semantic factors alone can be eliminated if the claim of exclusivity is made at the same time. Thus, a view that the lexicon is organizable solely in terms of a network of associative meaning (i.e. a thesaurus) can be said to be too restricted a view. Any partial-list approach, e.g., one that might be developed in response to Kiparsky's suggestion, in which certain morphemically related words are to be listed in the lexicon, has to have a way of relating them, or it must be shown that the underlying relationship among the variants is not productive. In other words, a list approach, full or partial, is not necessarily eliminated, if some way of relating items is postulated. Thus, Venneman's full word-list approach goes through; rules are retained to organize the lexicon and to analyze new forms. Hsieh's position is difficult to assess experimentally, since if subjects behave in accordance with the rule being tested the rule is said to be part of the internalized grammar, but not if they behave otherwise. This point is made not as a criticism of Hsieh's position, but to highlight its main concern, which is that generative rules are not to be proposed as part of a grammar until psychologically validated. Jackendoff's position shifts the issue from the psychological validation of generative rules to the psychological validation of lexical redundancy rules.

On the issue of the basis for the organization of lexical items, there are these possibilities: semantically based, syntactically based, phonologically based, or some combination of these. Evidence from aphasiology (for a review, see Baker, 1974), particularly the finding that one mode could malfunction without concommittant disruption of the other two modes, suggests that all three are each available for use as a basis for organizing lexical items. The question is whether there is one type of feature which predominates. Weigl and Bierwisch (1970) and Whitaker (1971) have suggested that semantic features play a more prominent role. Whitaker wrote:

> The data does strongly suggest that the semantic relationships between words is neurologically more important or prominent than phonetic similarity and thus indirectly suggests that the semantically organized lexicon is a better (simpler) model to relate to the central language system. (p. 172)

The findings of the present study suggest a kind of organization that is mainly phonological, partly syntactic, but certainly mot semantic. Do these findings, then, constitute counterevidence to Whitaker's conclusion? Let us first examine whether Whitaker's conclusion is warranted.

Whitaker's conclusion that semantic features play a more prominent role in lexical organization Was based on data of a sort typified by the following protocol:

> In attempting to read the sentence David wakes up early, R.H. [the patient] said:

> 44
> VF And what does he do early?
> R.H. To go to bed, early in the morning, very early in morning to go to sleep.
>
> The confusion here is go to sleep, go to bed with actual verb wakes up, clearly a similar semantic field showing a confusion of some of the semantic features but not others (Whitaker, 1971, p. 175).

Underlying Whitaker's conclusion is the assumption that if items in the lexicon are semantically organized, confusion would be in terms of semantically related words, as exemplified in the above protocol. The type of responses Whitaker would expect, if organization is phonological, is of a sort where the troublesome phrase wakes up would elicit phonetically related words, e.g., rakes up, makes up, takes up, etc.

The fact, however, that there is a type of speech disorder in which semantic knowledge stays intact does not constitute conlusive evidence for the primacy of semantic features. Other types of disorders may lead to a kind of performance that focuses on phonological features more than any other features, and, in a similar vein, this is not necessarily evidence for the primacy of phonological features, either. For example, a common speech "disturbance" normal speakers experience is the so-called "tip of the tongue" phenomenon (see, for example, Brown & McNeil, 1964). Speakers every now and then forget a word, usually a name, but they 'know' some features of the sound of the word, and a typical strategy in trying to retrieve the target word is to go through words that sound like the word to be recalled -- an indication that similar sounding words are somehow similarly indexed and referenced in the mental lexicon.

The kind of lexical organization revealed in aphasic and other speech disorder data may depend, not only on the type of disorder, but also on the manner in which data is accessed. As Baker (1974, p. 80) suggested, if semantic accessing is used, semantically relevant data will be elicited. The point can also be made that if accessing is phonological, phonologically relevant data will be elicited. Baker's experiments employed phonologically based elicitation technique (e.g., "Give as many words as you can think of that begin with the first sound of the word "wake".) and the present study focused on phonological knowledge; hence, both these studies elicited data that provide some information on phonological organization of lexical items.

The problem of the basis of organization in short term memory for words has also received some attention. Shulman (1972) wrote:

> The analysis of error data has frequently been used as a means of determining the nature of encoded information in short-term memory (STM). Although some investigators have found reliable evidence of semantically based errors in STM (e.g., Klein, 1970), the most common pattern of results has been that phonemic information is a more potent source of systematic errors than is semantic information (Baddeley, 1966; Kintsch & Buschke, 1969). These error data have led to the inference that semantic information is stored far less efficiently in STM than phonemic information. (p. 22a)

Indeed, Shulman's study, and an earlier one (Shulman, 1970), showed that subjects did in fact perform well on synonymy recognition test, given a data elicitation method that forced subjects to use semantic encoding.

In summary, it can be said that a list approach to the lexicon (partial or full, very likely a word-list, presumably not a morpheme-list) is a viable alternative to a fully generative approach. The contribution of the present study along this line is further experimental evidence that can be interpreted to mean that the lexicon _must_ be organized, and that certain items in the lexicon can be organized on the basis of phonological information, with very minimal syntactic information.

Implications to the Issue of Psychological Reality

On the basis of the demonstrated facilitated learnability of correct English alternations, the rash conclusion might be made that a particular formulation of the vowel alternation rule, for example, the Chomsky-Halle formulation, is psychologically real. Such a conclusion is presumptuous, to say the least. Findings of the sort reported in this thesis relate to the substance of the rule, and not to its form. Since only the substantive basis has been validated, other formulations of the rule that cover the same ground. Are in this sense likewise validated. What the study has established is that some such generalization must be part of an account of English if the account is to be adequate, since the generalization appears to have a psychological basis. The distinction being made here is important. The study has demonstrated a true statement about the English language, but not about how a particular theory of English should account for this fact about English.

The criticism may be made that this is nothing new, that grammarians have known all along that such vowel alternations obtain in English. In answer to this criticism, it may be said that what was known was the fact that certain adjective-noun pairs exhibit the vowel alternations that they exhibit, but it was not known that this pattern of alternations transcends existing words; that is, the "productivity" of the alternations cannot in principle be established, over and against the alternative that such alternations are simply memorized as parts of the "related" words, unless novel forms are shown to manifest the same alternations as real, already learned, forms.

One is reminded of a related, though not analogous, situation in early psycho-syntactic research. In the middle 1960's, chronometric studies yielded data that led to the acceptance of the view that transformations were psychologically real. As is now well-known, subsequent research and re-interpretation of the early data suggested that the observed effects are best attributed to semantic and constituent structural properties of sentences; in other words, to the substantive basis of sentences and not to rules that typologize sentences (for a concise review of these studies, cf. Baker, Prideaux, and Derwing, 1973).

The perfect fit between the experimental results and the predictions of Chomsky and Halle's and Schane's complexity measures based on marking conventions might lead some to make the unqualified conclusion that these conventions are psychologically real. Again, these conventions are to be regarded as real only in so far as they contain a description of certain processes that speakers genuinely employ in using their language. The conventions themselves are independent of these processes, and may take different forms. In fact, Terbeek (1973), after demonstrating the perceptual reality of certain markedness qualities of vowels, suggested that these qualities be treated as features, instead of calling in the theory of markedness to deal with them. The reasoning was that, if certain perceptually real vowel qualities, such as, for example, backness and roundedness, are regarded as features, then other vowel qualities which are demonstrable as likewise perceptually real, e.g., "tongue markedness" and "rounding markedness" (the markedness features Terbeek identified), should also be considered as features. In Chomsky and Halle's system, these later qualities are defined as abstractions in terms of the other features. As Terbeek pointed out, this "asymmetrical interpretation of perceptual dimensions is difficult to defend" (p. 676).

Now that vowel alternation in English has been demonstrated to be a variable in a mental activity, is vowel alternation then psychologically real? If psychological reality is defined in terms of demonstrable effect on the operation of certain mental processes, and there seems to be no reason why this definition cannot be made, then vowel alternation is psychologically real. Is it productive? In current linguistic literature, productivity is characterized somewhat narrowly and strictly as the creative use of rules, "creative" in the sense of combining forms to produce forms predicted by the rules. It is this definition of productivity that Steinberg and Krohn demonstrated to be an invalid characterization of English vowel alternation.

It is suggested that the following distinction be made: A productive rule is always psychologically real, but a psychologically real rule is not necessarily productive (given the definitions above). The threshold for psychological reality is lower than the threshold for productivity, so that it takes so much more for a psychologically real rule to be at the same time productive. "If a rule is highly productive, it readily applies to new items" (Krohn, 1972, p. 18; emphasis added). The distinction parallels somewhat that between "recognition threshold" and "recall threshold" in verbal learning theory (see Kausler, 1974, p. 8). Recall threshold is generally considered to be higher than recognition threshold. Recall of a form, which presupposes recognition, requires active production of the form, and, hence, it requires more thorough learning than recognition does.

This distinction provides a framework for explaining a number of things. First, the conflicting experimental results on vowel shift. The experiments by Steinberg and Krohn and by J. Ohala are productivity experiments, and this fact, in addition to the procedural shortcomings pointed out earlier, may have contributed to the negative results. The experiment by Moskowitz and the present experiment are psychological reality experiments, hence, the positive results. Second, the survey of ideas on the properties of a productive rule (see Chapter 1) led to the composite picture of a productive rule as exceptionless, surface phonological (as opposed to morphological), hence, phonetically motivated (as opposed to the arbitrariness of a morphological conditioning). This is a picture that does not fit English vowel alternation. If no claim of productivity is made for English vowel alternation, then the anomaly disappears.

And third, the distinction provides a somewhat informal explanation for the observation that, in normal conversation, vowel alternation is hardly ever adverted to in creative word play and word innovation, whereas English speakers everywhere often come up with the Correct plural forms of made-up words, for example, "glog[z]", "glit[s]", and "glitch[ɨs]". Compared to noun-stem pluralization, vowel alternation is unproductive. Vowel alternation has been weakened and is likely on the way out. Asked to give the adjective form equivalent of "serendipity", speakers often reply with "serendipous", seldom "serendipe". Even suffixation with "-ity", which is one environment for laxing, does not appear to be productive: the adjective "supine" elicits "supineness", not "sup[I]nity". There has in fact been no 'recent' addition to the closed set of English words that undergo vowel shift. But although vowel alternation is not productive (in the strict sense), there are enough extant forms in the language on which speakers can base a psychological generalization, hence, it is psychologically real.

In a more general sense, this notion of psychological reality/productivity -- for convenience, let us call it psychological productivity -- is an evaluation criterion that is capable not only of categorizing a rule as either psychologically productive or psychologically unproductive, but also of detecting gradations of psychological productivity. This is meant to parallel the notion of "scale of formal productivity" (see Krohn, 1972); the number of forms that obey particular rules varies from potentially infinite (e.g., noun pluralization) to a mere handful (e.g., g-deletion in "long"). One end of this scale can be labelled "Morphological" and the other end "Phonetic". The morphological end may represent "ungrammatized perceptual constraints" (Bever, 1970), which allow for graded judgment, and the phonetic end may represent "grammatized constraints", which accept only yes/no (i.e., productive/nonproductive) judgments. Thus, surface phonetic constraints, whose manifestations abound in any language, will be high in the scale of psychological productivity, whereas purely morphological rules, which apply on closed sets, will generally occupy lower rungs in this scale. The explanation for this gradation takes on functional considerations. Not only is there a high incidence of specific words that fall under the operation of surface phonetic constraints; in some of these cases, particularly in the case of noun pluralization in English, the facts are recapitulated in other processes: verbal inflection, possessive formation, is-contraction, has-contraction. In other words, the large functional load of this alternation has resulted in the "phoneticization" of what used to be apparently only accidentally related, morphologically conditioned sound changes. It is essentially this idea of functional load that formed the basis of Kiparsky's (1971) suggestion of a weaker version of his alternation condition, a version that "allow[s] absolute neutralization where the internal evidence includes several phonological processes of the language" (p. 590).

The idea that psychological productivity is graded (in J. R. Ross' terminology, "squishy") follows naturally if linguistic structures are treated as internal psychological structures; in other words, as Anderson (1972) pointed out, that grammatical structures are built from the basic units imposed by an inductive learning process (p. 412). Psychological structures are learned, and if learning is statistical and additive, as Anderson claimed it is (p. 422), then "an internal structure may have a 'strength' proportional to the amount of relevant experience" (p. 422). And certainly English speakers have far greater experience with the sibilant alternation than with the vowel alternations.

By way of a summary, this section has proposed a distinction between psychological reality and productivity in which psychological reality is treated as a necessary condition for productivity, but not vice versa. A scale of psychological productivity has been suggested, which may be used to evaluate morpho-phonological processes for external validity. Phonetic processes are expected to rate high in this scale, and purely morphological processes will rate low in this scale. The vowel alternation rules, which are morphological, gravitate towards the lower end; vowel alternation is still psychologically real, although not strong enough to exhibit high productivity.

A Note on 'Simplicity' and Learnability

In a recent paper, Braine (1974) wrote:

> Current generative theory predicts that the fewer the feature differences between such alternants, the easier should be the learning of the underlying rules: e.g., an alternation /u/ --> /ʌ/ should be more difficult to learn than /u/ --> /U/ or /u/ --> /o/, because rule simplicity depends on the number of features mentioned. My third hypothesis predicts that feature difference, per se, should not affect learning. Experimental test of these predictions should be possible with miniature artificial languages, or artificial extensions of natural languages. (p. 296)

We will examine the relevance of the results of the present experiment to the point Braine has raised in the above quotation. The experimentally determined learnability hierarchy of the alternations will be correlated with formally determined simplicity measures. The initial problem is to quantify the complexity of the alternations.

How does one quantify the complexity of say, the alternation ay/I? An obvious procedure is to locate the sounds in a vowel space and to measure the distance between them in terms of the number of features in which they differ. There are two problems associated with this task. First, what vowel space (i.e., feature system) to use, and, second, what sounds (underlying or surface) to locate in the vowel space. In what follows, the Chomsky-Halle system is assumed. Two distance metrics are proposed below.

Consider, first, a metric that focuses on surface shapes and ignores underlying sources. This surface alternation metric counts the number of features in which the vowels of the surface alternations differ. The features included are those related to tongue height, position, and rounding. Since the crucial vowels of the base words in this paradigm are all tensed and diphthongized, tensing and the glides will be ignored in the count. In this procedure, the alternation ay/I has the complexity index 3 (i.e., *a* and *I* differ in the features front, low, and high). The rest of the alternations have the following complexity indeces: iy/ɛ = 1, ey/æ = 1, aw/ʌ = 1, ow/ɔ = 1. It will be remembered that the learnability ranking of the alternations are (from highest to lowest) ow/ɔ, ay/I, iy/ɛ, ey/æ, and aw/ʌ. The learnability ranking does not support the formal complexity index.

A second system of quantifying the formal complexity of the alternations measures the distances of the alternate surface sounds from the postulated common underlying vowel form, and adding these two values together. For example, the postulated underlying vowel of the surface alternation ay/I is *i* (tensing is ignored). The distance between *i* and *a* is 3, and the distance between *i* and *I* is zero, giving a total complexity index of 3 for this alternation. The complexity indeces of the other alternations are iy/ɛ = 1, ey/æ = 1, aw/ʌ = 5, ow/ɔ = 1. The complexity index of the fourth alternation, aw/ʌ, has been computed as follows: The distance between underlying *u* and surface *a* is 3 (high, low, round), and the distance between underlying *u* and surface *ʌ* is 2 (high, round), giving a total complexity of 5. The learnability ranking and the formal complexity index of this alternation match, but no other formal predictions are supported by the experimental results. The formal complexity index predicts that the alternation ay/I should be next hardest to learn, a prediction that is not confirmed by the experimental results.

In the above computations of the complexity indices, the vowel feature specifications used was that of Chomsky and Halle's. Of course, the results may be different in other systems, and given other counting procedures. In so far as the results of the present experiment are concerned, Braine's expectation that formal simplicity and learnability of the alternations may not be isomorphic is supported.

Addendum: Other Explanations of the Results

The expperimental results have been explained in terms of linguistic habits or knowledge. Two other possible explanations are briefly mentioned below.

One possible source of the differential learnability of the alternations is frequency of occurrence of the alternations. If learnability is a function of frequency of occurrence, obviously the incorrect, non-occurring alternations would be disfavored. However, this cannot offer a complete explanation, as three of the 20 incorrect alternations (ay/ɔ, iy/æ, ey/ɔ) induced more learning than the correct alternation aw/ʌ. This writer does not discount the possibility that frequency of occurrence may have contributed to the total variation; however, it would require another major study to isolate the effect on learnability of frequency in this kind of a learning task.

An interesting possible explanation of the results has been suggested by J. Gray (personal communication). He suggested that the greater learnability of the correct English alternations may have been brought about by a strategy that makes use of orthographic representation of the alternations. Notice that of the five correct English alternations, all but one typically have the same letter representation in written English, and none of the incorrect alternations exhibits this property. Thus, ay/I is represented by the letter "i"; iy/ɛ by the letter "e"; ey/æ by the letter "a", ow/ɔ by the letter "o", and only the alternation aw/ʌ has different letter representations ("ou"/"u"). Now, if the subjects transformed the auditory signals into some orthographic representations, it may be easier to remember the pairs with the same spelling representation. If subjects used this strategy, the results are irrelevant to the major issue dealt with in this thesis. The suggestion is important and interesting, and merits some discussion.

The plausibility of Gray's suggestion depends on the assumption that subjects in fact made use of a mediational stage whereby the auditory signals were transformed into their orthographic representations (this possibility has been suggested by J. Ohala, 1974b), and that there was agreement among the subjects in just the right way of representing the sounds. It would be easy to find out if subjects would agree on the letter representations; they can be asked to write in conventional English spelling the stimulus words. But the results of this little test would be of little relvance to the main issue if the employment of an orthographic mediational stage cannot be demonstrated.

There is a piece of evidence against the possibility of a *purely* orthographic approach to the learning of the alternations: the fact that the alternation aw/ʌ, behaves like a true member of the class of the learnable correct English alternations. Thus, of the 25 alternations, it is ranked eighth in learnability; of the 125 possible stimulus by required response by actual response combinations, it is also ranked eighth. The point is that if the strategy used was completely orthographic, there would be no reason for this alternation, which violates the orthographic uniqueness requirement, to be ranked any higher than many other alternations which also violate this requirement.

CHAPTER SIX

SUMMARY, AND A PROPOSED STUDY

Summary

In the introductory chapter, a review of the properties of what linguists consider to be productive rules led to the conclusion, essentially a prediction, that vowel alternation in English is not productive. Since in current usage productivity and psychological reality are used synonymously, the conclusion can be interpreted to be a claim of psychological *un*reality for vowel alternation in English. A claim of psychological reality for vowel alternation in English, however, is implicit in Chomsky and Halle (1968).

In the review chapter, it was pointed out that both of these conflicting claims are in fact supported by experimental results. Steinberg and Krohn (1975), using a word-formation productivity test, showed that vowel alternation in English is unproductive. Moskowitz (1973), on the other hand, using a concept formation procedure, observed that vowel alternation is psychologically real.

Because of these conflicting experimental data, another experiment on the productivity/psychological reality of vowel alternation in English was deemed appropriate. A distinction was proposed between the psychological reality of the form of a rule and the psychological reality of its substance. It was made clear that the present experiment tested only for the latter type of psychological reality.

The logic of the present experiment is as follows: Assume that a linguistic construct is psychologically real if it is a variable in mental processing. Therefore, if correct English alternations are learned faster and better than incorrect English alternations, and if they interfere in the learning of incorrect English alternations, then knowledge of vowel alternation has been employed, and, thus, it is psychologically real.

A two-part learning experiment was conducted to assess the learnability and transferability of knowledge of correct English alternations, with non-English alternations serving as control conditions. The alternations were embedded in English-sounding nonsense word pairs, and these word pairs were presented aurally, using the study-test technique, as paired-associates to be learned. Forty university students served as subjects, half of whom did a task equivalent to forming words through front formation, and the other half did a task equivalent to forming words through back formation.

The results show that correct English alternations as a class induced better and faster learning than incorrect English alternations in both front and back formation tasks. However, only one subject learned to criterion, a fact that precluded treatment of the transfer scores as transfer scores. The transfer experiment was thus treated as a replication of the acquisition experiment, with comparable results.

The findings were discussed in terms of the notion of positive and negative interference. The greater learnability of items exemplifying correct English alternations was attributed to the facilitative effect (positive proactive interference) of knowledge of English morphology. It was suggested that this knowledge was the major source of negative interference in the learning of the non-English alternations. A second source of negative interference was suggested, which was interference due to the primacy of the peripheral or "tongue marked" vowels.

The applicability of this knowledge of the alternations to meaningless forms suggests that the items in the lexicon are in fact relatable on the basis of phonological information, supplemented only by the barest of syntactic information. An explanation for this linguistic knowledge was sought for, and it was pointed out that while both associationistic and organizational models of verbal learning and behavior can account for the observed greater frequency of correct English responses, neither model serves a useful role in shedding any light as to the initial learning of these alternations.

The findings strongly suggest that the *substantive* basis of vowel alternation in English is psychologically real. It was made clear that the experimental data do not say anything at all about the reality of the *form* of the vowel alternation rules. Since the content of VAR has been found to be psychologically real, it was also concluded that any rule that captures this aspect of English is, in just this sense, valid.

In an attempt to reconcile the seemingly incompatible experimental results on vowel alternation in English, a distinction between psychological reality and productivity was suggested, whereby psychological reality is treated as a necessary, though not sufficient, condition for productivity, and that a psychologically real rule need not be productive. Thus, vowel alternation may not be productive, but it is psychologically real. A scale of "psychological productivity" was proposed as an external, evaluative criterion for rules, and it was suggested that surface phonetic processes rate high on this scale and non-phonetically conditioned alternations rate low on this scale.

A Proposed Study

Now that there is evidence for the substantive validity of the alternations, how this generalization is to be represented, i.e., the form of the rule, may be tested. One source of disagreement as to how the vowel shift rule is to be written concerns the matter of whether it is a two-stage process, as in Chomsky and Halle's version, or a one-stage process, as suggested by Wang (1968). In the two-stage formulation, only two of the vowels, i.e., [i, u], show changes at each stage of the derivation. The other vowels make use of either the first stage or the second stage. The two-stage derivation is illustrated below, where the input strings have already undergone tensing and diphthongization:

	Divine	profound
Input	div[iy]ne	prof[uw]nd
Vowel Shift, Stage 1	div[ey]ne	prof[ow]nd
Vowel Shift, Stage 2	div[æy]ne	prof[ɔw]nd
Output, after Backness adj.	div[ay]ne	prof[aw]nd

Since the difference between the output of Stage 2 and the final output is not relevant to height, the final output will be treated, for the sake of simplicity, as the output of Stage 2. Thus the stages in the derivation are: [i] --> [e] --> [a] and [u] --> [o] --> [a]. What is the reality of the intermediate stage [e]/[o]?

Using a slightly different set of features (i.e., [mid] replacing [low]), Wang (1968) formulated a shift rule comparable to Chomsky and Halle's except that Wang's is a one-stage process. In this process, [i] directly changes to [æ], and [u] to the low back (in Wang's terms, nonhigh, nonmid) [ɒ]. However, Wang's system is not perfect; two of the vowels have to undergo what amounts to an intermediate stage. Thus, [æ] --> [ɛ], and [ɒ] --> [ɔ], but the correct surface forms are [e] and [o], respectively. Wang proposed an independently motivated redundancy rule that raises mid vowels to high. Applied to the output of the shift rule, this rule will give the desired results. But notice that Wang's proposal, in effect, amounts to a two-stage shift derivation for two of the vowels (not the same two vowels in the Chomsky-Halle formulation). The difference in the two approaches is that Wang views each of the 'two' stages of the shift to be manifestations of two different processes, one phonological and the other morpheme structural. Chomsky and Halle, on the other hand, view the shifts as two stages of the same phonological process. Thus, a comparable question may be asked of Wang's proposal: What is the reality of the intermediate stage [ɛ]/[ɔ]?

One of the most widely used paradigms in verbal learning research is serial learning. It is ideal for assessing the degree of intralist organization. If the items in a list have a high degree of organization, learning and recall of the items is facilitated. A common type of organization is chained association of the elements. For example, the letters of the alphabet can be recited back easily when they are to be given in order. The sequence 1,2,3 is easier to remember than, say, 2,1,3, or 2,3,1, as a consequence of the obvious sequential organization of the first series. In a sequential organization, the elements are chained together as a string of stimulus and response units, with a non-terminal element serving as response to the previous element and stimulus to the next one

The idea of chained association of items in serial learning can serve as a basis for an experimental hypothesis to test for the validity of intermediate stages in a derivation. If there is any reality to these stages, and in fact, in the ordering of these stages, the output of the derivational stages arranged in order should be easier to learn or to remember than the same output elements arranged differently. If Stage 1 feeds Stage 2 and Stage 2 feeds Stage 3, then the output of these stages in this order should facilitate learning, and any other ordering of the output should retard learning. This appears to be a reasonable interpretation of the psychological implications of stages in a formal derivation.

If the sort of experiments envisioned here is to be feasible at all, the input to the rule being tested and the output of each derivational stage should be presentable in some way to speakers. Fortunately, the input to vowel shift (after tensing and diphthongization) as well as the output of the intermediate stages are utterable. They can therefore be presented aurally to the subjects. Thus, the problem of whether the intended sound has been heard as such by the subject does not enter into the picture. (If the underlying segment is an abstract segment, there is no way to find out if subjects actually perceive the segment as such, or whether a process of synthesis has mediated. For example, /ɛ, ɔ/ are presumably abstract for Nupe speakers. If nonsense forms with such segments are presented, the question arises as to whether Nupe subjects would perceive them as such or whether these segments would be automatically re-interpreted as, say, [e, o].)

In the proposed experiment, nonsense forms such as, for example, perh[iy]n-perh[ey]n-perh[ay]n, could be presented, to test for the Chomsky-Halle derivation. To test for the Wang derivation, nonsense forms such as perh[æy]n-perh[ɛy]n-perh[ey]n could be presented. The subjects will be 'probed' on their memory for the forms. Given perh[iy]n as probe, the expectation is that perh[ey]n, i.e., the output of the next stage of the derivation in the Chomsky-Halle formulation, should constitute majority of the subjects' primary responses, if the hypothesis of fit between psychological derivation and the Chomsky-Halle formal derivation is to be maintained. Similarly, if the probe is perh[ey]n, the expected response is perh[ay]n. To control for the well-observed primacy effect (which is that the first and last elements of a sequence are learned best, see Jahnke, 1963; Underwood, 1963), each stimulus set will be presented in all possible orders, thus: 1,2,3; 1,3,2; 2,1,3; etc. The patterns of results could be positive for both Chomsky and Halle and Wang, or negative for both, or positive for one and negative for the other. The last possibility should be most interesting. Positive results for both will be difficult to explain, whereas negative results for both may mean the unreality of intermediate derivational stages of any sort.

An experiment of the sort being proposed is a logical extension of the present study. The present study provides a confirmation of the reality of the substance of the vowel shift rule. The question of the reality of its form may now be addressed with some justification.

REFERENCES

Anderson, L. Explanation, abstractness, and language learning. In P. Peranteau, J. Levi & G. Phares (Eds.), Papers from the eighth regional meeting of the Chicago Linguistic Society, 1972.

Anisfeld, M. Psychological evidence for an intermediate stage in a morphological derivation. Journal of Verbal Learning and Verbal Behavior, 1968, 8, 191-195.

Baddeley, A. Short-term memory for word-sequences as a function of acoustic, semantic, and formal simplicity. Quarterly journal of Experimental Psychology, 1966, 18, 362-365.

Baker, L. The lexicon: some psycholinguistic evidence. Working papers in phonetics, No. 26. University of California at Los Angeles, 1974.

Baker, W. & Prideaux, G. Word category effects on paired-associate learning: a critique of Glanzer's results. Paper presented at the meeting of the Canadian Linguistic Association, Edmonton, Alberta, May 1975.

Baker, W., Prideaux, G. & Derwing, B. Grammatical propertieis of sentences as a basis for concept formation. Journal of Psycholinguistic Research, 1973, 2, 201-220.

Berko, J. The child's learning of English morphology. Word, 1958, 14, 150-157.

Bever, T. A survey of some recent work in psycholinguistics. In W.J. Plath (Ed.), Specification and utilization of a transformational grammar. Yorktown Heights, New York: IBM Corp., 1968.

Bever, T. The cognitive basis for linguistic structures. In J. Hayes (Ed.), Cognition and the development of language. New York: John Wiley & Sons, Inc., 1970.

Bever, T., Fodor, J. & Garrett, M. A formal limitation of associationism. In T. Dixon and D. Horton (Eds.), *Verbal behavior and general behavior theory*. Englewood Cliffs, N.J.: Prentice-Hall, 1968.

Bloomfield, L. *Language*. New York: Holt, Rinehart, and Winston, 1933.

Bower, G. Organization factors in memory. *Cognitive Psychology*, 1970, *1*, 18-46.

Braine, M. Two models of the internalization of grammars. In D. Slobin (Ed.), *The ontogenesis of grammar*. New York: Academic Press, 1971.

Braine, M. On what might constitute learnable phonology. *Language*, 1974, *50*, 270-299.

Brame, M. On the abstractness of phonology: Maltese ¿. In M. Brame (Ed.), *Contributions to generative phonology*, Austin, Texas: University of Texas Press, 1972.

Brown, R. & McNeil, D. The tip of the tongue phenomenon. *Journal of Verbal Learning and Verbal Behavior*, 1966, *5*, 325-337.

Bugarski, R. Writing systems and phonological insights. *Papers from the sixth regional meeting of the Chicago Linguistic Society*, 1970.

Cearley, A. The only phonological rule ordering principle. In A. Bruck, R. Fox, & M. La Galy (Eds.), *Papers from the parasession on natural phonology*. Chicago, Illinois: Chicago Linguistic Society, 1974.

Chomsky, C. Reading, writing, and phonology. *Harvard Educational Review*, 1970, *40*, 287-309.

Chomsky, N. *Syntactic structures*. The Hague: Mouton, 1957.

Chomsky, N. Review of B. F. Skinner, *Verbal behavior*. *Language*, 1959, *35*, 26-58. Reprinted in L. Jakobovits & M. Miron (Eds.), *Readings in the psychology of language*. Englewood Cliffs, New Jersey: Prentice-Hall, 1967. (Page references to the latter.)

Chomsky, N. The general properties of language. In C. Millikan & F. Darley (Eds.), *Brain mechanisms underlying speech and language*. New York: Grune and Stratton, 1967.

Chomsky, N. *Language and mind*. New York: Harcourt, Brace & World, 1968.

Chomsky, N. Linguistics and philosophy. In S. Hook (Ed.), *Language and philosophy*. New York: New York University Press; 1969.

Chomsky, N. Remarks on nominalization. In R. Jacobs and P. Rosenbaum (Eds.), *Readings in English transformational grammar*. Waltham, Mass.: Blaisdell, 1970.

Chomsky, N. & Halle, M. *The sound pattern of English*. New York: Harper & Row, 1968.

Cofer, C. Verbal learning. In M. Marx (Ed.), *Learning: Processes*. New York: The Macmillan Company, 1969.

Cofer, C. Properties of verbal materials and verbal learning. Chap. 19 in J. Kling & L. Riggs (Eds.), *Woodworth and Schlosberg's Experimental Psychology* (3rd ed.). New York: Holt, Rinehart, & Winston, 1971.

Coltheart, M. & Geffen, G. Grammar and memory I: phonological similarity and proactive interference. *Cognitive Psychology*, 1970, *1*, 215-224.

Crothers, J. On the abstractness controversy. *Project on linguistic analysis*, No 12, University of California at Berkeley, 1971.

Cutting, J. Aspects of phonological fusion. *Journal of Experimental Psychology*, 1975, *104*, 105-120.

D'Anglejan, A., Lambert, W., Tucker, G. & Greenberg, J. Psychological correlates of the French sound system. *Perception & Psychophysics*, 1971, *9*(3B), 356-357.

Day, R. Fusion in dichotic listening. Unpublished doctoral dissertation, Stanford University, 1968.

Day, R. Temporal order of perception of a reversible phoneme cluster. *Haskins Laboratories status report on speech research*, 1970, *24*, 47-56.

Day, R. & Cutting, J. Levels of processing in speech perception. Paper presented at the 10th annual meeting of the Psychonomic Society, San antonio, Texas, November 1970.

Denes, P. On the statistics of spoken English. *The Journal of the Acoustical Society of America*, 1963, *35*, 892-904.

Derwing, B. *Transformational grammar as a theory of language acquisition*. New York & London: Cambridge University Press, 1973.

Derwing, B. What kind of rules can children learn? In press.

Dewey, G. *Relative frequency of English speech sounds*. Cambridge, Mass.: Harvard University Press, 1923.

Esper, E. A technique for the experimental investigation of associative interference in artificial linguistic material. *Language Monographs*, 1925, No. 1.

Fillenbaum, S. Psycholinguistics. *Annual Review of Psychology*, 1971, *22*, 251-308.

Fodor, J. & Bever, T. The psychological reality of linguistic segments. *Journal of Verbal Learning and Verbal Behavior*, 1965, *4*, 414-420.

Fodor, J., Bever, T. Garrett, M. *The psychology of language*. New York: McGraw-Hill Book Company, 1974.

Fodor, J. & Garrett, M. Some reflections on competence and performance. In J. Lyons & R. Wales (Eds.), *Psycholinguistic papers*. Edinburgh: Edinburgh University Press, 1966.

Fromkin, V. Simplicity is a complicated question. *Working papers in phonetics*, No. 17. University of California at Los Angeles, 1971.

Fromkin, V. (Ed.). *Speech errors as linguistic evidence*. The Hague: Mouton, 1973.

Glanzer, M. Grammatical category: a rote learning and word association analysis. *Journal of Verbal Learning and Verbal Behavior*, 1962, *1*, 31-41.

Glanzer, M. Psycholinguistics and verbal learning. In K. Salzinger & S. Salzinger (Eds.), *Research in verbal behavior and some neurophysiological implications*. New York: Academic Press, 1967,

Goldstein, L. & Lackner, J. Alterations of the phonetic coding of speech sounds during repetition. *Cognition*, 1973, *2*, 279-297.

Greenberg, J. *Language universals*. The Hague: Mouton & Co., 1966.

Greenberg, J. & Jenkins, J. Studies in the psychological correlates of the sound system of American English. *Word*, 1964, *20*, 157-177.

Gough, P. Experimental psycholinguistics. In W. Dingwall (Ed.), *A survey of linguistic science*. Linguistics Program, University of Maryland, 1971.

Hale, K. The passive and ergative in language change: the Australian case. In S. Wurn and D. Laycock (Eds.), *Pacific linguistics studies in honour of Arthur Capell*. Canberra: Australian National University, 1970.

Halle, M. Stress rules in English: a new version. *Linguistic Inquiry*, 1973, *4*, 452-464.

Halle, M. & Keyser, S. *English stress*. New York: Harper & Row, 1971.

Hankamer, J. Analogical rules in syntax. In P. Peranteau, J. Levi & G. Phares (Eds.), *Papers from the eighth regional meeting of the Chicago Linguistic Society*, 1972.

Harms, R. How abstract is Nupe? *Language*, 1973, *49*, 439-446.

Hsieh, H-I. The psychological reality of tone sandhi rules in Taiwanese. Papers from the sixth regional meeting of the Chicago Linguistic Society, Chicago, Illinois, 1970.

Hsieh, H-I. Lexical diffusion: evidence from child language acquisition. Glossa, 1972, 6, 89-104.

Hsieh, H-I. How generative is phonology? In E.F.K. Koerner (Ed.), Current issues in linguistic theory. Amsterdam: John Benjamin B.V., 1975.

Hsieh, H-I. On the unreality of some phonological rules. Lingua, 1976, 38, 1-19.

Hudson, G. On the representation of non-productive alternation. In J. Anderson & C. Jones (Eds.), Historical linguistics II. New York: American Elsevier Publishing Company, Inc., 1974.

Hyman, L. How concrete is phonology? Language, 1970, 46, 58-76.

Hyman, L. Nupe three years later. Language, 1973, 49, 447-452.

Irwin, O. Infant speech: development of vowel sounds. Journal of Speech and Hearing Disorders, 1948, 13, 31-34.

Jackendoff, R. Morphological and semantic regularities in the lexicon. Language, 1975, 51, 639-671.

Jackson, F. An experimental study of English word stress. Proceedings of the 48th annual meeting of the Linguistic Society of America, 1973, p. 76. (Summary)

Jahnke, J. Serial position effects in immediate serial recall. Journal of Verbal Learning and Verbal Behavior, 1963, 2, 284-287.

Jakobson, R. Child language, aphasia, and phonological universals. The Hague: Mouton, 1968.

Jenkins, J., Foss, B. & Greenberg, J. Phonological distinctive features as cues in learning. Journal of Experimental Psychology, 1968, 77, 200-205.

Johnson, N. A model of sentence generation. Paper presented at the meeting of the American Psychological Association, Los Angeles, California, 1964.

Johnson-Laird, P. Experimental psycholinguistics. Annual Review of Psychology, 1974, 25, 135-160.

Kausler, D. Psychology of verbal learning and memory. New York: Academic Press, 1974.

Kenyon, J. & Knott, T. A pronouncing dictionary of American English. Springfield, Mass.: Merriam, 1953.

Kimble, G. Hilgard and Marquis' conditioning and learning. New York: Appleton-Century-Crofts, 1961.

Kintsch, W. & Buschke, H. Homophones and synonyms in short-term memory. Journal of Expperimental Psychology, 1969, 80, 404-407.

Kiparsky, P. Linguistic universals and linguistic change. In E. Bach & R. Harms (Eds.), Universals in linguistic theory. New York: Holt, Rinehart, & Winston, 1968.

Kiparsky, P. Historical linguistics. In W. Dingwall (Ed.), A survey of linguistic science. Linguistics Program, University of Maryland, 1971.

Kiparsky, P. Metrics and morphophonemics in the Rigveda. In M. Brame (Ed.), Contributions to generative phonology. Austin: University of Texas Press, 1972.

Kiparsky, P. Phonological representations. In O. Fujimura (Ed.), Three dimensions of linguistic theory. Tokyo: TEC Company, Ltd., 1973. (a)

Kiparsky, P. Productivity in phonology. In M. Kenstowicz & C. Kisseberth (Eds.), Issues in phonological theory. The Hague: Mouton, 1973. (b)

Kiparsky, P. Remarks on analogical change. In J. Anderson & C. Jones (Eds.), *Historical linguistics II*. New York: American Elsevier Publishing, 1974. (a)

Kiparsky, P. On the evaluation metric. In A. Brock, R. Fox & M. La Galy (Eds.), *Papers from the parasession on natural phonology*. Chicago, Illinois: Chicago, Linguistic Socity, 1974. (b)

Kisseberth, C. On the abstractness of phonology: the evidence from Yawelmani. *Papers in Linguistics*, 1969, *1*, 248-282.

Klatt, D. Structure of confusions in short-term memory between English consonants. *Journal of the Acoustical Society of America*, 1968, *44*, 401-407.

Klein, G. Temporal changes in acoustic and semantic confusion effects. *Journal of Experimental Psychology*, 1970, *86*, 236-240.

Koefoed, G. On formal and functional explanation: some notes on Kiparsky's "Explanation in phonology". In J. Anderson & C. Jones (Eds.), *Proceedings of the 1st international conference on historical linguistics*. Dordrecht: Reidel, 1975.

Kramsky, J. The frequency of occurrence of vowel phonemes in languages possessing vowel systems of identical structure. In L. Dolezel, P. Sgall & J. Vacheck (Eds.), *Prague studies in mathematical linguistics*. University of Alabama Press, 1966.

Krohn, R. The vowel shift rule and its productivity. *Language Sciences*, 1972, *20*, 17-18.

Krohn, R. Is there a rule of absolute neutralization in Nupe? *Glossa*, 1976, *9*, 139-148.

Labov, W. Methodology. In W. Dingwal (Ed.), *A survey of linguistic science*. Linguistics Program, University of Maryland, 1971.

Lackner, J. & Goldstein, L. The psychological representation of speech sounds. *Quarterly Journal of Experimental Psychology*, 1975, *27*, 173-185.

Ladefoged, P. An alternative set of vowel shift rules. *Working papers in phonetics*, No. 17. University of California at Los Angeles, 1971.

Ladefoged, P. & Broadbent, D. Perception of sequence in auditory events. *Quarterly Journal of Experimental Psychology*, 1960, *13*, 162-170.

Ladefoged, P. & Fromkin, V. Experiments on competence and performance. *IEEE Transactions on Audio- and Electro-Acoustics*, 1968, *AU-16*, 130-136.

Lightner, T. The role of derivational morphology in generative grammar. *Language*, 1975, *51*, 617-638.

Liljencrants, J. & Lindblom, B. Numerical simulation of vowel quality systems: the role of perceptual contrast. *Language*, 1972, *48*, 839-862.

Maher, J. Etymology and generative phonology in traditional lexicon. *General Linguistics*, 1971, *11*, 71-98.

Mandler, G. Verbal learning. In G. Mandler, P. Mussen, N. Kogan, & M. Wallach (Eds.), *New Directions in Psychology* (Vol. 3). New York: Holt, Rinehart & Winston, 1967.

McGuire, W. A multi-process model for paired-associate learning. *Journal of Experimental Psychology*, 1961, *62*, 335-347.

Merriam-Websters' Third New International Dictionary. Springfield, Mass.: Merriam, 1966.

Miller, G. *Language and communication*. New York: McGraw-Hill Book Company, Inc., 1951.

Miller, G. The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 1956, *63*, 81-97.

Miller, G., Galanter, E. & Pribram, K. *Plans and the structure of behavior*. New York: Holt, Rinehart, and Winston, 1960.

Moorehead, D. Processing of phonological sequences by young children and adults. *Papers and reports on child language development*, No. 1. Stanford University, 1970.

Moskowitz, B.A. On the status of vowel shift in English. In T. Moore (Ed.), *Cognitive development and the acquisition of language*. New York: Academic Press, 1973.

Neisser, L. *Cognitive psychology*. New York: Appleton-Century-Crofts, 1967.

Nessly, L. Experimental phonology and English Stress. Paper presented at the meeting of the Linguistic Society of American, Ann Arbor, Michigan, Summer, 1973. (a)

Nessly, L. A^{1}thelsta^{3}n meets A^{1}thelsta^{0}n. Paper presented at the 48th annual meeting of the Linguistic Society of America, San Diego, California, December 1973. (b)

Nie, N., Hull, C., Jenkins, J., Steinbrenner, K. & Bent, D. *Statistical package for the social sciences* (2nd ed.). New york: McGraw-Hill, Inc., 1975.

Norman, D. *Memory and attention: an introduction to human information processing*. New York: Wiley and Sons, 1969.

Ohala, J. Aspects of the control and production of speech. *Working papers in phonetics*, No. 15. University of California at Los Angeles, 1970.

Ohala, J. Experimental historical phonology. In J. Anderson & C. Jones (Eds.). *Historical Linguistics II*. New York: American Elsevier Publishing, 1974. (a)

Ohala, J. On the design of phonological experiments. Unpublished manuscript, University of Claifornia at Berkeley, 1974. (b)

Ohala, J. & Ohala, M. Testing hypothesis regarding the psychological manifestation of morpheme structure conditions. Paper presented at the meeting of the Linguistic Society of America, San Francisco, California, December 1975.

Ohala, M. The abstractness controversy: experimental input from Hindi. Language, 1974, 50, 225-235.

Olofsson, A. Latin rules in English phonology? Linguistic Inquiry, 1974, 5, 320-322.

Oxford English Dictionary. Oxford: Clarendon Press, 1933.

Paul, H. Principles of the history of language. Translated by H. Strong. London: Longmans and Green, 1891.

Prytulak, L. Natural language mediation. Cognitive Psychology, 1971, 2, 247-254.

Random House Dictionary of the English Language. New York: Random House, 1966.

Rommetveit, R. Words, meanings, and messages. New York: Academic Press, 1968.

Ross, J. A re-analysis of English word stress. In M. Brame (Ed.), Contributions to generative phonology. Austin, Texas: University of Texas Press, 1972.

Saltarelli, M. Italian qua neo-Latin. In J. Casagrande & B. Saciuk (Eds.), Generative studies in Romance Languages. Rowley, Mass.: Newbury House, 1972.

Saltarelli, M. A Latin conspiracy. To appear in a volume in honor of Archibald A. Hill.

Schane, S. [back] and [round]. In S. Anderson & P. Kiparsky (Eds.), A festschrift for Morris Halle. New York: Holt, Rinehart, and Winston, 1973.

Schane, S. Noncyclic English word stress. Unpublished manuscript, University of California at San Diego, 1974.

Schane, S., Tranel, B. & Lane, H. On the psychological reality of a natural rule of syllable structure. Cognition, 1974, 3, 351-358.

Schmerling, S. A stress mess. Studies in the Linguistic

Sciences, 1971, *1*, 52-66.

Schnitzer, M. *Generative phonology -- evidence from aphasia*. University Park, Pennsylvania: Administrative Committee on Research, Pennsylvania State University, 1972.

Schnitzer, M. Aphasiological evidence for five linguistic hypothesis. *Language*, 1974, *50*, 300-315.

Scholes, R. Performance aspects of readjustment rules. In P. Peranteau, J. Levi & G. phares (Eds.), *Papers from the eighth regional meeting of the Chicago Linguistic Society*, 1972.

Settera, G. A non-cyclical approach to stress placement. *Glossa*, 1974, *8*, 83-108.

Sherzer, J. Talking backwards in Cuna: the sociological reality of phonological descriptions. *Southwestern Journal of Anthropology*, 1970, *26*, 343-353.

Shibatani, M. The role of surface phonetic constraints in generative phonology. *Language*, 1973, *49*, 87-106.

Shulman, H. Encoding and retention of semantic and phonemic information in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1970, *9*, 499-508.

Shulman, H. Semantic confusion errors in short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 1972, *11*, 221-227.

Skousen, R. On capturing regularities. In P. Peranteau, J. Levi & G. phares (Eds.), *Papers from the eighth regional meeting of the Chicago Linguistic Society*, Chicago, Illinois, 1972.

Sloat, C. Stress in English. *Glossa*, 1974, *8*, 121-139.

Steinberg, D. Psychological aspects of Chomsky's competence-performance distinction. *Working papers in linguistics*, Vol. 2, No. 2, Department of Linguistics, University of Hawaii, 1970.

Steinberg, D. & Krohn. R. The psychological reality of Chomsky & Halle's Vowel Shift Rule. In E. Koerner (Ed.), *The transformational paradigm and modern linguistic theory*. Amsterdam: John Benjamins B.V., 1975.

Stockwell, R. Problems in the interpretation of the great English vowel shift. In M. Smith (Ed.), *Studies in linguistics in honor of George L. Trager*. The Hague: Mouton, 1972.

Terbeek, D. Six dimensions of vowel quality. In C. Corum, T. Smith-Stark & A. Weiser (Eds.), *Papers from the ninth regional meeting of the Chicago Linguistic Society*. Chicago, Illinois: Chicago Linguistic Society, 1973.

Thorndike, E. & Lorge, I. *The teacher's word book of 30,000 words*. New York: Teacher's College Press, 1944.

Underwood, B. Stimulus selection in verbal learning. In C. Cofer and B. Musgrave (Eds.), *Verbal behavior and learning: problems and processes*. New York: McGraw-Hill, 1963.

Venneman, T. Words and syllables in natural generative grammar. In A. Bruck, R. Fox, & M. La Galy (Eds.), *Papers from the parasession on natural phonology*. Chicago, Illinois: Chicago Linguistic Society, 1974.

Voelker, C. Technique for a phonetic frequency distribution count in formal American speech. *Arch. neerl. Phon. exp.*, 1935, *11*, 69-72.

Wang, M. What stops a sound change? In J. Casagrande & B. Saciuk (Eds.), *Generative studies in Romance languages*. Rowley, Mass.: Newbury House, 1972.

Wang, M. Productivity in generative phonology. Unpublished doctoral dissertation, University of Texas at Austin, 1974.

Wang, W. Vowel features, paired variables, and the English vowel shift. *Language*, 1968, *44*, 695-708.

Wang, W. Competing changes as a cause of residue. *Language*, 1969, *45*, 9-25.

Warren, R. Illusory changes in repeated words: differences between young adults and the aged. *American Journal of Psychology*, 1961, *74*, 536-516.

Warren, R. & Gregory, R. An auditory analogue of the visual reversible figure. *American Journal of Psychology*., 1958, *71*, 612-613.

Warren, R. & Warren, R. Auditory illusions and confusions. *Scientific American*, 1970, *233*, 30-36.

Weigl, E. & Bierwisch, M. Neuropsychology and linguistics: topics of common research. *Foundations of language*, 1970, *6*, 1-18.

Whitaker, H. *On the representation of language in the human brain*. Edmonton, Alberta: Linguistic Research, Inc., 1971.

Wickelgren, W. Distinctive features and errors in short-term memory for English vowels. *Journal of the Acoustical Society of America*, 1965, *38*, 583-588.

Wickelgren, W. Distinctive features and errors in short-term memory for English consonants. *Journal of the Acoustical Society of America*, 1966, *39*, 388-398.

Wickelgren, W. Context-sensitive coding in speech recognition, articulation, and development. In K. Leibovic (Ed.), *Information processing in the nervous system*. New York: Springer-verlag, 1969. (a)

Wickelgren, W. Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, 1969, *76*, 1-15. (b)

Zimmer, K. Psychological correlates of some Turkish Morpheme structure rules. *Language*, 1969, *45*, 309-321.

Zwicky, A. Note on a phonological hierarchy in English. In R. Stockwell & R. Macaulay (Eds.), *Linguistic change and generative theory*. Bloomington, Indiana: Indiana University Press, 1972.

APPENDIX A
STIMULUS MATERIALS

	Acquisition	Transfer
"divine/divinity"-Group		
ay/I	subk[ay]pe-subk[I]pity	disk[ay]pe-disk[I]pity
ay/ɛ	malp[ay]de-malp[ɛ]dity	purp[ay]de-purp[ɛ]dity
ay/æ	purg[ay]te-purg[æ]tity	eng[ay]te-eng[æ]tity
ay/ʌ	surf[ay]me-surf[ʌ]mity	malf[ay]me-malf[ʌ]mity
ay/ɔ	disr[ay]ne-disr[ɔ]nity	exr[ay]ne-exr[ɔ]nity
"serene/serenity"-Group		
iy/I	surk[iy]te-surk[I]tity	subd[iy]te-subd[I]tity
iy/ɛ	malm[iy]de-malm[ɛ]dity	purm[iy]de-purm[ɛ]dity
iy/æ	disf[iy]me-disf[æ]mity	malf[iy]me-malf[æ]mity
iy/ʌ	del[iy]ne-del[ʌ]nity	disl[iy]ne-disl[ʌ]nity
iy/ɔ	ent[iy]pe-ent[ɔ]pity	mist[iy]pe-mist[ɔ]pity
"profane/profanity"-Group		
ey/I	prel[ey]be-prel[I]bity	surl[ey]be-surl[I]bity
ey/ɛ	subn[ey]ze-subn[ɛ]zity	maln[ey]ze-maln[ɛ]zity
ey/æ	deb[ey]me-deb[æ]]mity	disb[ey]me-disb[æ]mity
ey/ʌ	surt[ey]ce-surt[ʌ]city	ext[ey]ce-ext[ʌ]city
ey/ɔ	disd[ey]pe-disd[ɔ]pity	ent[ey]pe-ent[ɔ]pity
"profound/profundity"-Group		
aw/I	misn[aw]d-misn[I]dity	pren[aw]d-pren[I]dity
aw/ɛ	surs[aw]n-surs[ɛ]nity	diss[aw]n-diss[ɛ]nity
aw/æ	pren[aw]se-pren[æ]sity	den[aw]se-den[æ]sity
aw/ʌ	perc[aw]t-perc[ʌ]tity	surc[aw]t-surc[ʌ]tity
aw/ɔ	imp[aw]ze-imp[ɔ]zity	malp[aw]ze-malp[ɔ zity
"verbose/verbosity"-Group		
ow/I	end[ow]ne-end[I]nity	misd[ow]ne-misd[I]nity
ow/ɛ	emp[ow]be-emp[ɛ]bity	surp[ow]be-surp[ɛ]bity
ow/æ	purh[ow]de-purh[æ]dity	preh[ow]de-preh[æ]dity
ow/ʌ	disl[ow]pe-disl[ʌ]pity	perl[ow]pe-perl[ʌ]pity
ow/ɔ	exg[ow]me-exg[ɔ]mity	eng[ow]me-eng[ɔ]mity

Note. The underlined word pairs are correct English exemplars; all other word pairs served as control pairs. The relevant sounds are enclosed in square brackets and written in phonetic spelling; the irrelevant parts are written in standard English spelling.

APPENDIX B
SUBJECT INSTRUCTIONS

Acquisition Test

This is a study to find out more about memory. Pairs of words will be played on a tape recorder and your task is to try to remember these pairs of words. In this test, we are not trying to find out how good your memory is; we are interested in the characteristics of people's memories. We want to find which pairs are easy to learn, and which pairs are hard to learn. Probably the pairs that give you the most trouble are the ones that are difficult for other people, too. Knowing which pairs are generally most difficult to remember will allow us to look for the reasons, and, perhaps, to find out more about how mmemory works.

The experiment is divided into two parts. The first part is a familiarization pretest, in which you will familiarize yourself with the pronunciation of the words you will later try to learn. The second part consists of two tasks. In the first task, you will learn (memorize) pairs of words; there are 25 pairs to be learnned. In the second task, you will be asked to apply what you have learned to a new set of words.

The learning task consists of two phases: the study phase and the test phase. In the study phase, you will hear all 25 pairs, presented one after another. This is the time when you will try to remember the pairings of the words. A short gap separates the words of a pair, and a longer gap separates pairs. When all the pairs have been presented, you will hear a short high tone. This is the signal that that particular study period is over and that a test phase will immediately follow.

In the test phase, you will hear only the first word of each pair, and you will respond with the other word. When all 25 pairs have been tested, you will hear a short high tone, which signals the end of that test and the start of the next study phase. You will go into this cycle of study-test periods ten times, or until you have learned all the pairs, whichever is first.

In the early test phases, you may find it difficult to recall the correct response words. Whenever you are stuck for an answer, you are encouraged to make an intelligent guess.

If you have any questions so far, please ask them. If none, please continue reading.

We will now go into the familiarization pretest. The words that you will study in this experiment are not real English words. They are made-up or invented words, but they have been constructed in such a way that they sound like real English words. As you hear a word, please repeat it. (Please ask the experimenter to start the tape recorder.) The words that you just heard will be paired with the words that you will now hear. Pronounce the words as you hear them. (Please ask the experimenter to start the tape recorder.)

It has probably become evident to you that the pairs are intended to illustrate a pattern of noun-adjective alternation in English where the noun form ends in "-ity", and the word pairs have the same general meaning. For example, the pair _plausible_-_plausibility_,* as used in the sentences:

Jack offered a PLAUSIBLE resolution of the dilemma.

The PLAUSIBILITY of his solution attracted the attention of environmentalists.

As a possible aid to learning the nonsense word pairs, you may think of them as adjective and noun pairs exemplifying the same grammatical relation exhibited In pairs such as _plausible_-_plausibility_, _possible_-_possibility_, _tenable_-_tenability_, And many others.

The order of the word pairs will be varied in every presentation.

If you have no other questions, we will now begin the first study phase.

Please pronounce your answers as clearly as possible. Your responses will be recorded on tape.

If you are ready for the first study phase, say "Ready".**

---

*This is the order of the illustrative words in the front formation condition, i.e., base word first and derived word second. Appropriate adjustments were made in the instructions to subjects who belonged to the back formation condition.

**In the first two trials, E identified the task involved, thus: " "This is a study phase. You will hear both words of each pair. You will not make a response." Or, "This is a test. You will hear only the first word. You are to identify the word paired with it."

Transfer Test

In this test, you will work on a new set of words. Unlike in the first task, however, you will not hear the complete pair, but only the test word of each pair. You are asked to supply the correct partner of the test word. The purpose of this test is to find out if you can apply whatever you have learned in the learning phase to a new set of words of similar construction. We expect you to rely on your knowledge of the nature of the pairing of words in the first task to enable you to supply the correct pairs. You are encouraged to try a response even if you are not certain of it.

There are 25 test words; a short silence separates words. You will make your response within this short interval. The test words will be presented to you until you are able to give their respective partners or until three blocks of presentation have been completed, whichever is first.

Please ask all the questions you may have in mind about the procedure of this part of the experiment.

If you are ready for this test, say "Ready".

APPENDIX C
TABLES

TABLE 1

FREQUENCY DATA, SUMMED OVER SUBJECTS AND TRIALS

Front Formation							Back Formation						
S	RR	Actual Response					S	RR	Actual Response				
		I	ɛ	æ	ʌ	ɔ			ay	iy	ey	aw	ow
	I	154	31	6	3	2		ay	117	38	25	2	6
	ɛ	83	92	4	4	12		iy	82	60	14	11	5
ay/	æ	67	27	70	13	12	I/	ey	31	18	90	6	26
	ʌ	32	16	60	83	3		aw	47	28	20	63	4
	ɔ	28	18	2	4	131		ow	50	25	18	30	50
	I	75	68	5	38	2		ay	67	68	25	5	11
	ɛ	28	142	4	2	10		iy	33	95	17	5	10
iy/	æ	19	25	113	35	4	ɛ/	ey	33	11	99	19	13
	ʌ	52	81	4	48	3		aw	40	16	32	51	24
	ɔ	67	61	8	7	53		ow	42	56	20	17	29
	I	99	25	53	4	12		ey	39	10	48	36	36
	ɛ	24	45	88	2	31		iy	43	40	45	9	21
ey/	æ	6	9	168	4	7	æ/	ey	21	6	118	6	25
	ʌ	14	12	70	92	8		aw	18	11	74	54	21
	ɔ	25	17	27	22	101		ow	30	10	56	14	52
	I	64	26	11	34	48		ay	68	35	28	9	20
	ɛ	11	84	13	25	61		iy	28	70	38	8	21
aw/	æ	22	23	77	8	62	ʌ/	ey	31	13	76	18	25
	ʌ	38	14	3	100	34		aw	55	24	17	41	17
	ɔ	9	15	13	12	135		ow	21	25	25	13	82
	I	45	30	7	15	91		ay	95	14	18	16	34
	ɛ	25	54	7	12	95		iy	45	37	45	9	35
ow/	æ	14	12	81	4	72	ɔ/	ey	27	14	72	13	47
	ʌ	14	19	2	100	58		aw	29	10	28	59	44
	ɔ	2	3	4	4	185		ow	14	12	18	21	115

Notes. S = Stimulus Type, RR = Required Response
The underlined entries are Correct English scores.

TABLE 2

SUMMARY TABLE, STIMULUS x ACTUAL RESPONSE

Front Formation						Back Formation					
Stim.	Actual Response					Stim.	Actual Response				
	I	ɛ	æ	ʌ	ɔ		ay	iy	ey	aw	ow
ay	364	184	142	107	160	I	327	169	167	112	91
iy	241	377	134	130	72	ɛ	215	246	193	97	87
ey	168	108	406	124	159	æ	151	77	341	119	155
aw	144	162	117	179	340	ʌ	203	167	184	89	165
ow	100	118	101	135	501	ɔ	210	87	181	118	275

Note. Entries along the main diagonal are Correct English scores.

TABLE 3

SUMMARY TABLE, REQUIRED RESPONSE x ACTUAL RESPONSE

Front Formation						Back Formation					
RR	Actual Response					RR	Actual Response				
	I	ɛ	æ	ʌ	ɔ		ay	ay	ey	aw	ow
I	437	180	82	94	155	ay	386	165	144	68	107
ɛ	171	417	116	45	209	iy	231	302	159	42	92
æ	128	96	509	64	157	ey	143	62	455	62	136
ʌ	15	142	139	423	106	aw	189	89	171	268	110
ɔ	131	114	54	49	605	ow	157	128	137	95	328
	1017	949	900	675	1232		1106	746	1066	335	773

Note. Entries along the main diagonal are Correct Test scores.

TABLE 4

OVERALL ANOVA RESULTS

Source	df	MS	F
A (Formation)	1	28.94	25.84**
B (Trial)	9	3.33	26.71***
C(A) (C=Stimulus)	8	0.49	8.19***
D(A) (D=Required Resp.)	8	0.16	0.86
E(A) (E=Actual Resp.)	8	41.94	5.81***
S(A) (S=Subject)	6	1.12	
AB	9	0.58	4.65***
BC(A)	72	0.08	1.12
BD(A)	72	0.10	1.52*
CD(A)	32	0.23	2.15**
BE(A)	72	1.60	1.76***
CE(A)	32	56.03	16.94***
DE(A)	32	136.30	33.59***
BS(A)	54	0.12	
CS(A)	24	0.06	
DS(A)	24	0.19	
ES(A)	24	7.22	
BCD(A)	288	0.09	1.62***
BCE(A)	288	0.93	1.72***
BDE(A)	288	2.36	4.25***
CDE(A)	128	7.41	5.11***
BCS(A)	216	0.07	
BDS(A)	216	0.06	
CDS(A)	96	0.10	
BES(A)	216	0.91	
CES(A)	96	3.31	
DES(A)	96	4.06	
BCDE(A)	1152	0.67	1.41***
BCDS(A)	864	0.06	
BCES(A)	864	0.54	
BDES(A)	864	0.55	
CDES(A)	384	1.45	
BCDES(A)	3456	0.48	

*$p < .05$ **$p < .01$ ***$p < .001$

TABLE 5

RESULTS OF POST-HOC COMPARISONS AMONG SELECTED CATEGORIES: SCHEFFE'S COMPARISON VALUES AND SIGNIFICANCE

Categories Compared		Means	Comparison Value
A. Correct English (CE) vs. Incorrect English (ICE)			
Front:	CE	3.74	
	ICE	0.84	2.90**
Back:	CE	2.68	
	ICE	0.78	1.90**
Overall:	CE	3.21	
	ICE	0.94	2.27**
D. Correct English (CE) vs Controls			
Front:			
1. ay/I:	CE	3.85	
	Controls	0.26	3.59**
2. iy/ɛ:	CE	3.57	
	Controls	0.26	3.29**
3. ey/æ:	CE	4.20	
	Controls	0.16	4.04**
4. aw/ʌ:	CE	2.50	
	Controls	0.56	1.94**
5. ow/ɔ:	CE	4.62	
	Controls	0.08	4.54**
Back:			
1. I/ay:	CE	2.95	
	Controls	0.44	2.51**
2. ɛ/iy:	CE	2.67	
	Controls	0.34	2.33**
3. æ/ey:	CE	3.15	
	Controls	0.27	2.88**
4. ʌ/aw:	CE	1.35	
	Controls	0.69	0.66 n.s.
5. ɔ/ow:	CE	3.30	
	Controls	0.32	2.98**
Overall:			
1. ay/I:	CE	3.40	
	Controls	0.35	3.05**
2. iy/ɛ:	CE	3.12	
	Controls	0.31	2.91**

(table continued on next page)

TABLE 5 (continued)
RESULTS OF POST-HOC COMPARISONS AMONG SELECTED CATEGORIES:
SCHEFFE'S COMPARISON VALUES AND SIGNIFICANCE

Categories Compared	Means	Comparison Value
3. ey/æ: CE	3.68	
Controls	0.22	3.46**
4. aw/ʌ: CE	1.30	
Controls	0.63	1.30 n.s.
5. ow/ɔ: CE	3.96	
Controls	0.20	3.76**
C. Correct Test (CT) vs. Incorrect Test (ICT)		
Front: CT	3.74	
ICT	0.84	2.90**
Back: CT	2.68	
ICT	0.78	1.90**
Overall: CT	3.21	
ICT	0.80	2.41**
D. CE and CT vs. ICE and CT (i.e., CE vs. ICE, with CT effect held constant)		
Front: CT and CE	3.74	
CT and ICE	2.05	1.69**
Back: CT and CE	2.68	
CT and ICE	1.76	0.92**
Overall: CT and CE	3.21	
CT and ICE	1.91	1.20**
E. CT and ICE vs. ICT and ICE (i.e., CT vs. ICT, with English effect held constant)		
Front: ICE and CT	2.05	
ICE and ICT	0.59	1.46**
Back: ICE and CT	1.76	
ICE and ICT	0.57	1.19**
Overall: ICE and CT	1.91	
ICE and ICT	0.58	1.33**
F. Peripheral (P) vs. Non-Peripheral Vowels (NP)		
Front: Peripheral [I, ɔ]	28.06	
NP [ɛ, æ, ʌ]	20.98	7.08**
Back: P [ay, ow]	22.50	
NP [iy, ey, aw]	20.22	2.28**
Overall: 'Peripheral'	25.28	
'Non-Peripheral'	20.60	4.68**

TABLE 6

ANOVA RESULTS, CORRECT TEST SCORES

Source	df	MS	F
Front Formation			
A (Stimulus)	4	15.34	2.12
B (Required Response)	4	63.79	8.80***
AB	16	92.32	12.74***
Subject x AB	475	7.25	
Back Formation			
A (Stimulus)	4	9.38	1.91
B (Required Response)	4	101.54	20.67***
AB	16	56.17	11.44***
Subject x AB	475	4.91	

***p < .001

TABLE 7

ANOVA RESULTS, FIRST CORRECT TRIAL SCORES

Source	df	MS	F
Front Formation			
A (Stimulus)	4	5.45	0.64
B (Required Response)	4	37.89	4.41**
AB	16	72.49	8.44***
Subject x AB	475	8.58	
Back Formation			
A (Stimulus)	4	19.04	1.44
B (Required Response)	4	90.82	6.85***
AB	16	27.44	2.07**
Subject x AB	475	13.25	

**p <. 01 ***p < .001

TABLE 8

ANOVA RESULTS, LAST ERROR TRIAL SCORES

Source	df	MS	F
Front Formation			
A (Stimulus)	4	26.09	2.62*
B (Required Response)	4	63.66	6.39***
AB	16	85.74	8.61***
Subject x AB	475	9.96	
Back Formation			
A (Stimulus)	4	5.22	0.69
B (Required Response)	4	23.52	3.13**
AB	16	26.45	3.52***
Subject x AB	475	7.51	

*p < .05 **p < .01 ***p < .001

TABLE 9

CORRECT ENGLISH SCORES AS PERCENT OF TOTAL ERROR SCORES

Front Formation						Back Formation					
Stim.	Actual Response					Stim.	Actual Response				
	I	ɛ	æ	ʌ	ɔ		ay	iy	ey	aw	ow
ay	-	81	56	29	54	I	-	62	32	45	66
iy	60	-	30	58	43	ɛ	67	-	35	45	27
ey	56	61	-	67	30	æ	64	60	-	52	57
aw	29	23	7	-	24	ʌ	55	35	51	-	47
ow	64	68	71	62	-	ɔ	41	27	49	54	-

TABLE 10

ANOVA RESULTS: TRIAL x STIMULUS x REQUIRED RESPONSE x CORRECT ENGLISH/CORRECT TEST

Source	df	MS	F
Front Formation			
A (Trial)	9	9.88)4b	10.39***
B (Stimulus)	4	40.75	42.82***
C (Required Response)	4	7.49	7.87***
D (Cor. Eng/Cor. Test)	1	159.05	167.15***
AB	36	.70	0.74
AC	36	.74	0.78
BC	16	92.23	96.93***
AD	9	18.57	19.53***
BD	4	35.56	37.38***
CD	4	36.59	38.46***
ABC	144	.75	0.79
ABD	36	1.25	1.32
ACD	36	1.07	1.13
BCD	16	20.52	21.57***
ABCD	144	1.21	1.28***
Subj x ABCD	1500	.95	
Back Formation			
A (Trial)	9	23.04	23.24***
B (Stimulus)	4	27.74	27.98***
C (Required Response)	4	9.58	9.66***
D (Cor. Eng/Cor. Test)	1	221.11	223.06***
AB	36	1.57	1.58*
AC	36	1.28	1.30
BC	16	46.13	46.54***
AD	9	11.20	11.30***
BD	4	24.06	24.27***
CD	4	16.55	16.69***
ABC	144	1.25	1.27
ABD	36	1.24	1.25
ACD	36	1.08	1.09
BCD	16	12.85	12.96***
ABCD	144	1.02	1.03***
Subj x ABCD	1500	.99	

TABLE 11

ERROR SCORES NOT ATTRIBUTABLE TO CORRECT ENGLISH, SUMMARIZED AS A FUNCTION OF STIMULUS (ROWS) AND ACTUAL RESPONSE (COLUMNS)

Front Formation						Back Formation					
Stim.	Actual Response					Stim.	Actual Response				
	I	ɛ	æ	ʌ	ɔ		ay	iy	ey	aw	ow
ay	-	112	29	60	64	I	-	109	77	49	41
iy	88	-	28	20	114	ɛ	148	-	94	46	58
ey	61	71	-	56	85	æ	112	37	-	65	103
aw	118	61	69	-	48	ʌ	135	95	108	-	83
ow	103	53	27	37	-	ɔ	115	50	109	59	-
	370	297	153	173	311		510	293	388	219	285

TABLE 12

ERROR SCORES NOT ATTRIBUTABLE TO CORRECT ENGLISH, SUMMARIZED AS A FUNCTION OF REQUIRED RESPONSE (ROW) AND ACTUAL RESPONSE (COLUMN)

Front Formation						Back Formation					
RR	Actual Response					RR	Actual Response				
	I	ɛ	æ	ʌ	ɔ		ay	iy	ey	aw	ow
ay		92	72	24	29	I	-	97	96	59	73
iy	166	-	21	82	19	ɛ	149	-	114	34	57
ey	69	63	-	32	58	æ	112	51	-	44	89
aw	80	78	40	-	205	ʌ	142	73	97	-	66
ow	55	64	20	35	-	ɔ	107	72	81	82	-
	370	297	153	173	311		510	293	388	219	285

TABLE 13

RESULTS OF DISCRIMINANT FUNCTION ANALYSIS OF CORRECT TEST SCORES, WITH SS AS VARIABLES AND ALTERNATIONS AS GROUPS

Front Formation

Discriminant Function	Eigenvalue	Relative Percentage	Canonical Correlation
1	18233.76	99.44	1.00
2	86.44	0.47	0.99
3	13.29	0.07	0.96
4	2.46	0.07	0.84

Function Derived	Wilk's Lambda	Chi-Square	df	Significance
0	0.0000	209.10	80	.000
1	0.0002	96.27	57	.001
3	0.0202	44.86	36	.148
3	0.2890	14.27	17	.648

Centroids of Groups in Reduced Space:

	Function 1	Function 2	Function 3	Function 4
Group [ɪ]	1.18	-1.13	0.07	-0.90
Group [ɛ]	-1.21	-0.90	1.04	0.51
Group [æ]	-0.48	-0.19	-1.77	0.37
Group [ʌ]	-0.62	1.36	0.23	1.06
Group [ɔ]	1.14	0.87	0.42	1.06

(table continued on next page)

TABLE 13 (continued)

RESULTS OF DISCRIMINANT FUNCTION ANALYSIS OF CORRECT TEST SCORES, WITH SS AS VARIABLES AND ALTERNATIONS AS GROUPS

Back Formation

Discriminant Function	Eigenvalue	Relative Percentage	Canonical Correlation
1	354.31	79.60	1.00
2	76.92	17.28	0.99
3	11.45	2.57	0.96
4	2.45	0.55	0.84

Function Derived	Wilk's Lambda	Chi-Square	df	Significance
0	0.00	160.87	80	.000
1	0.00	93.33	57	.002
2	0.02	43.24	36	.190
3	0.29	14.24	17	.650

Centroids of Groups in Reduced Space:

	Function 1	Function 2	Function 3	Function 4
Group [ay]	1.23	-1.35	0.04	-0.57
Group [iy]	0.67	0.75	-1.23	0.90
Group [ey]	-1.64	-0.91	-0.46	0.20
Group [aw]	-0.30	1.18	0.02	-1.28
Group [ow]	0.02	0.33	1.64	0.76

TABLE 14

ANOVA RESULTS, TRANSFER SCORES

Source	df	MS	F
Front Formation			
A (Stimulus)	4	0.03	0.05
B (Required Response)	4	0.02	0.04
C (Actual Response)	4	8.92	14.58***
AB	16	0.03	0.05
AC	16	15.12	24.70***
BC	16	24.43	39.91***
ABC	64	3.28	5.36***
Subject x ABC	2375	0.61	
Back Formation			
A (Stimulus)	4	0.31	0.49
B (Required Response)	4	0.12	0.18
C (Actual Response)	4	14.23	22.28***
AB	16	0.06	0.10
AC	16	10.48	16.40***
BC	16	9.26	14.49***
ABC	64	1.98	3.11***
Subject x ABC	2375	0.64	

***$p < .001$

APPENDIX D
FIGURES

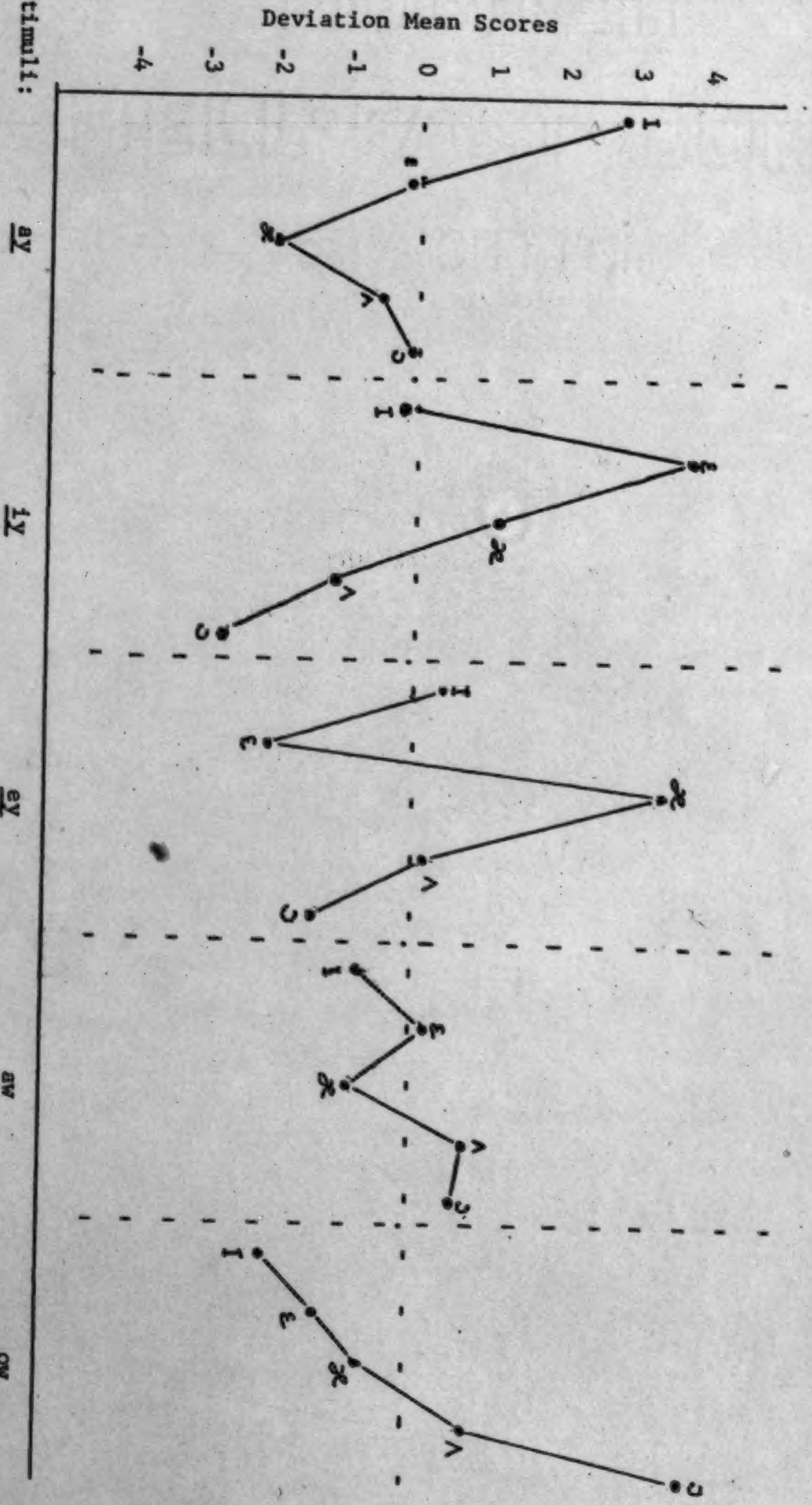

Figure 1A. Stimulus-Actual Response Interaction, Correct Test Scores, Front Formation

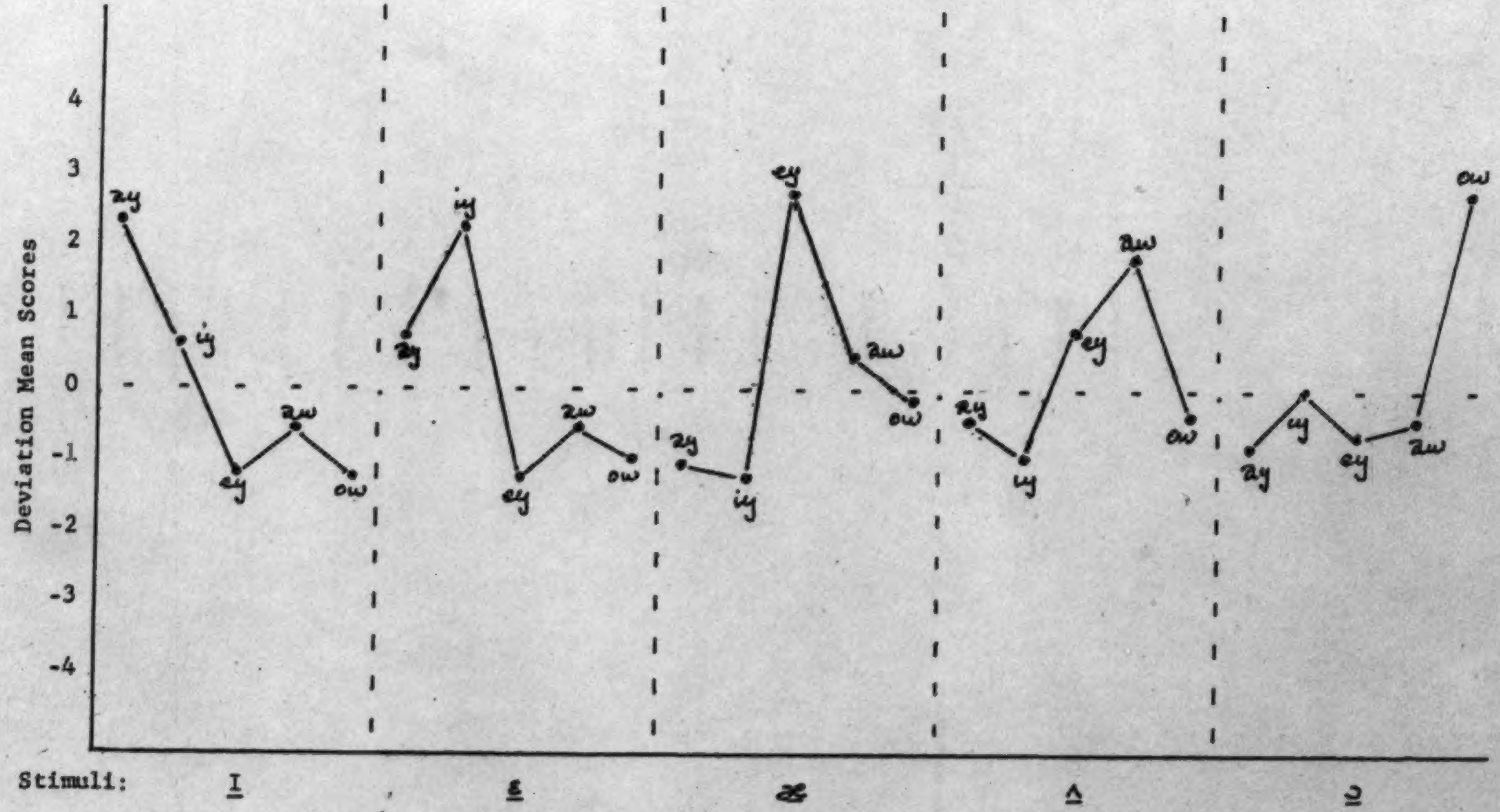

Figure 1B. Stimulus-Actual Response Interaction, Correct Test Scores, Back Formation

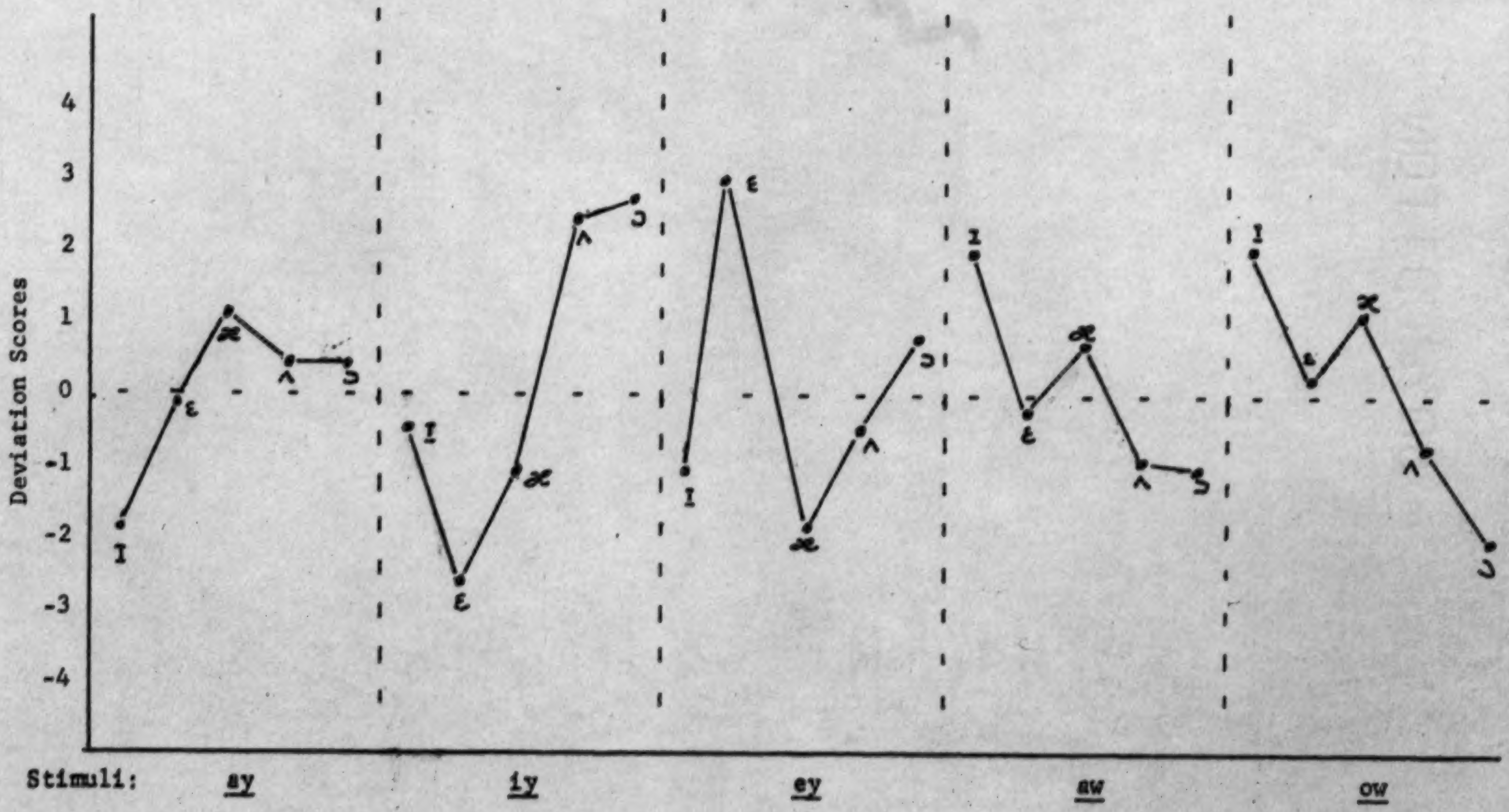

Figure 2A. Stimulus-Actual Response Interaction, First Correct Trial Scores, Acquisition Test, Front Formation

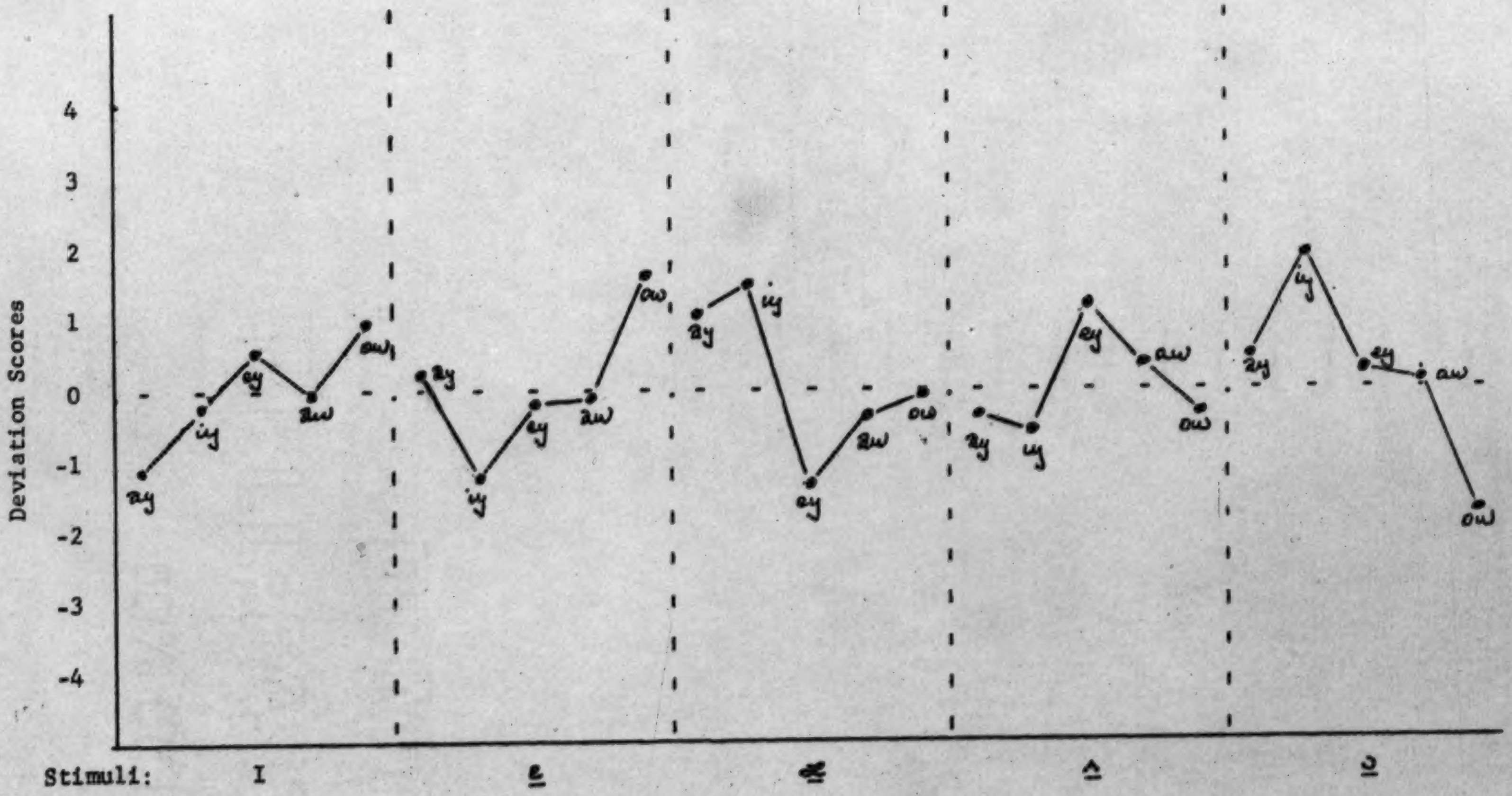

Figure 2B. Stimulus-Actual Response Interaction, First Correct Trial Scores, Acquisition Test, Back Formation

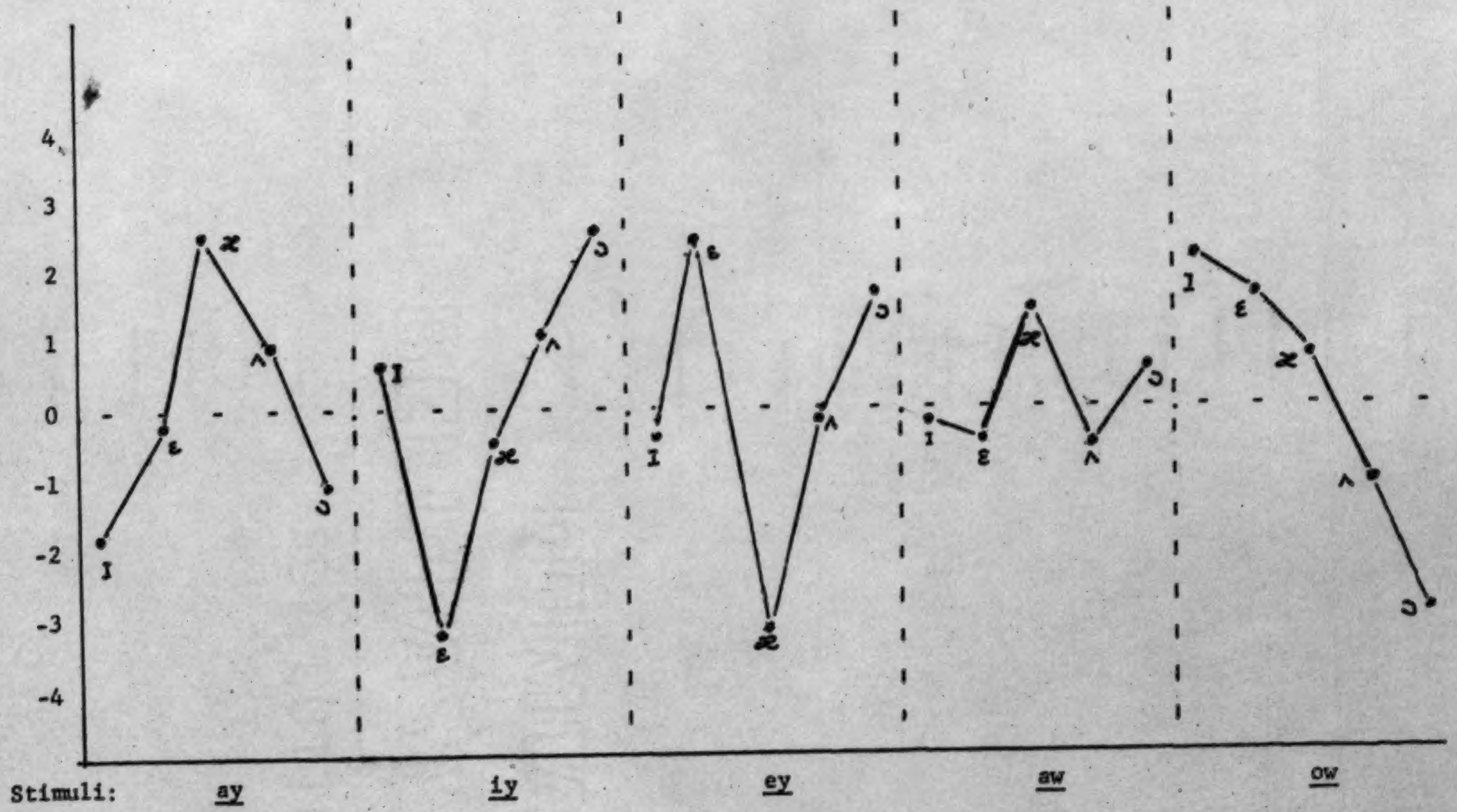

Figure 3A. Stimulus-Actual Response Interaction, Last Error Trial Scores, Front Formation

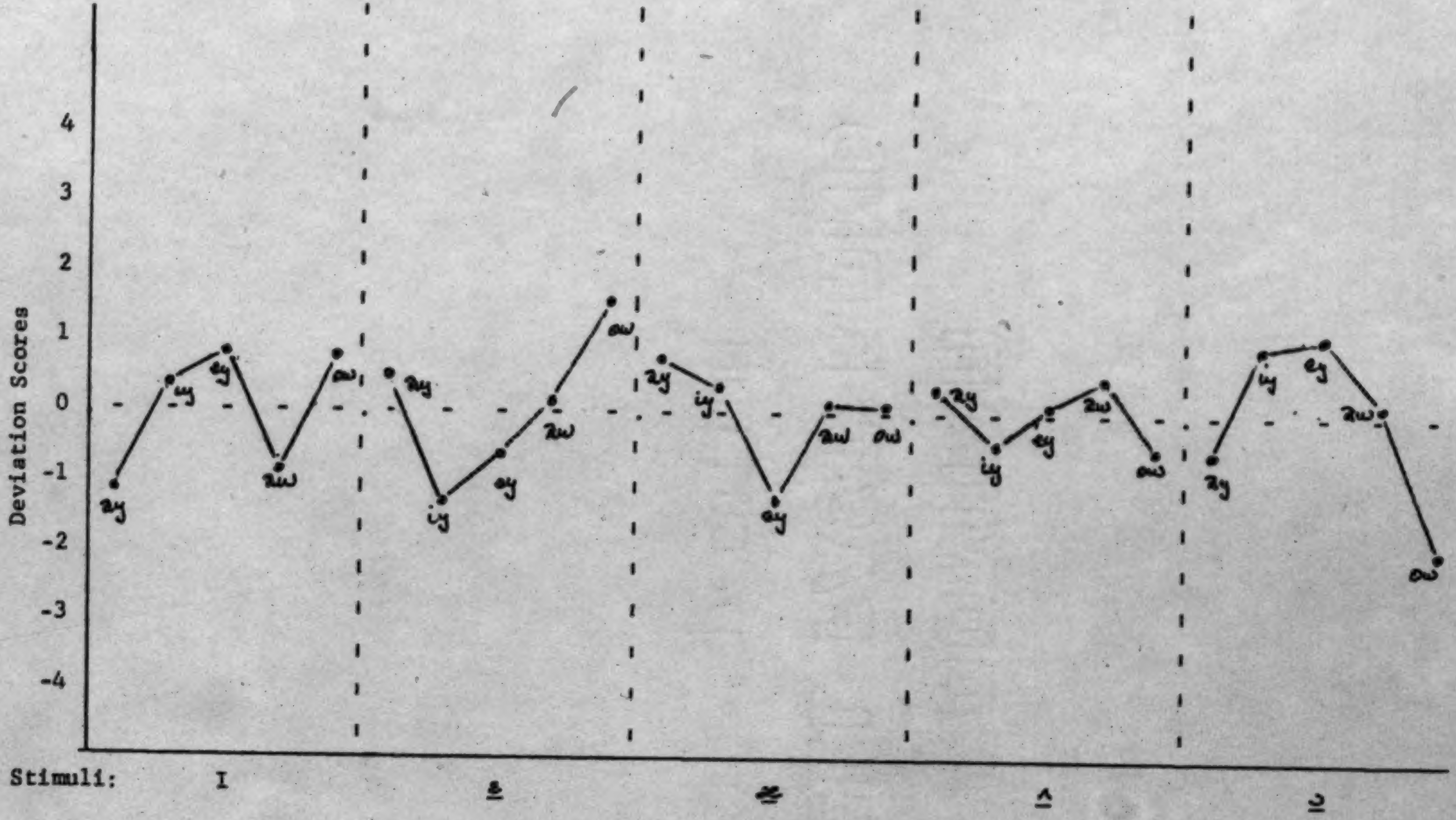

Figure 3B. Stimulus-Actual Response Interaction, Last Error Trial Scores, Back Formation

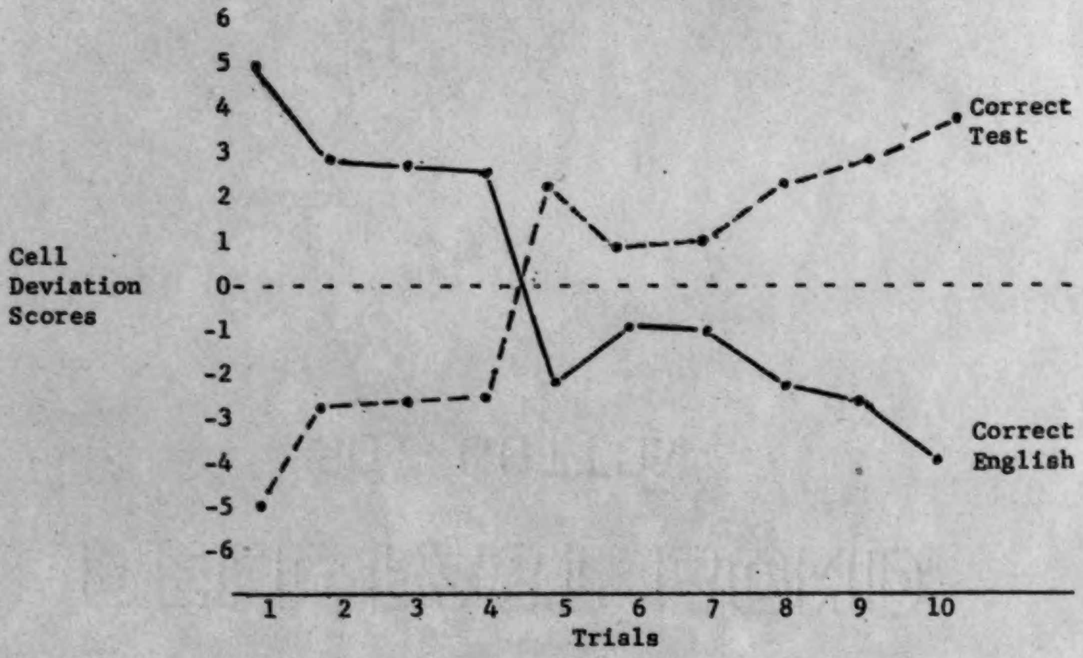

Figure 4A. Cell Deviation Scores of the Categories Correct Test and Correct English, Front Formation

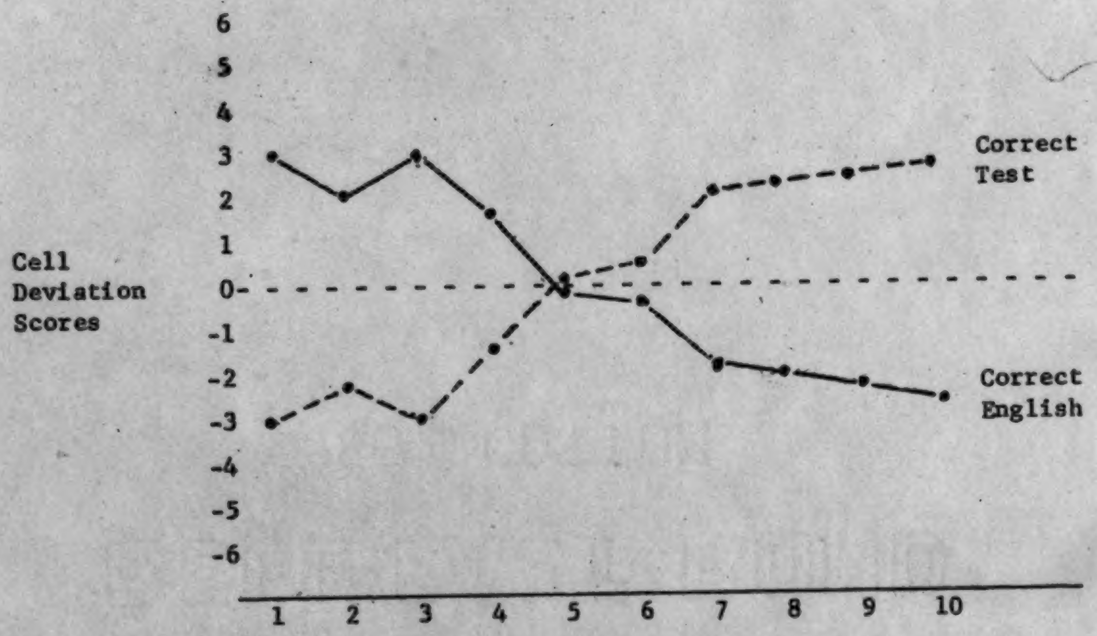

Figure 4B. Cell Deviation Scores of the Categories Correct Test and Correct English, Back Formation

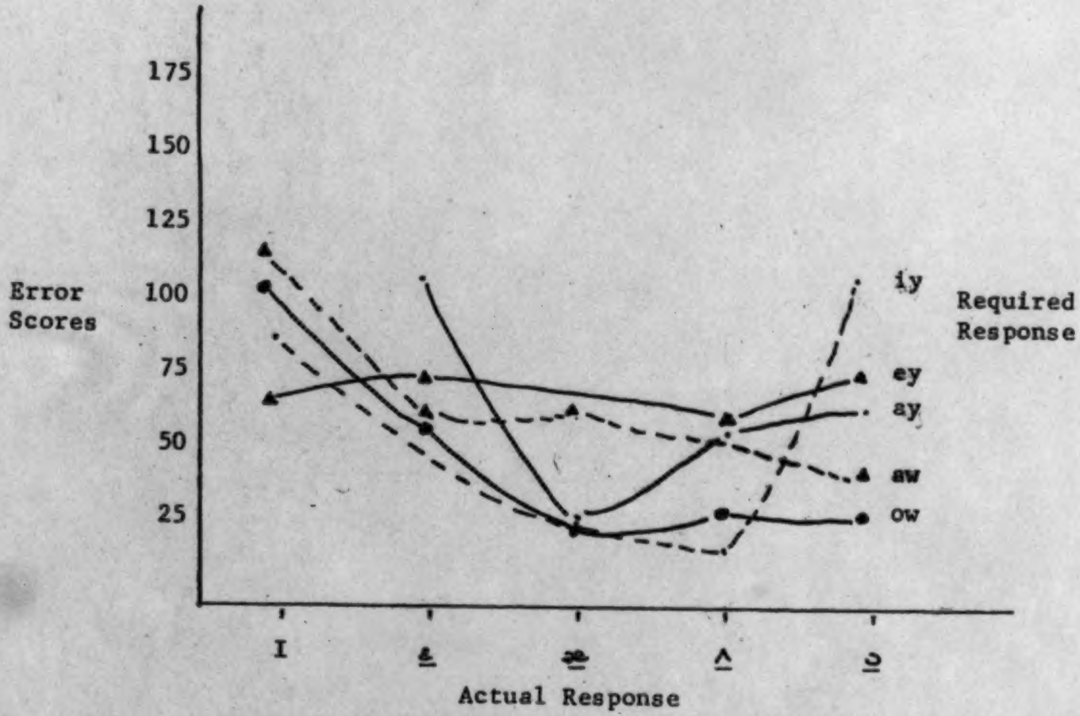

Figure 5. Front Formation Error Scores not Attributable to Correct English

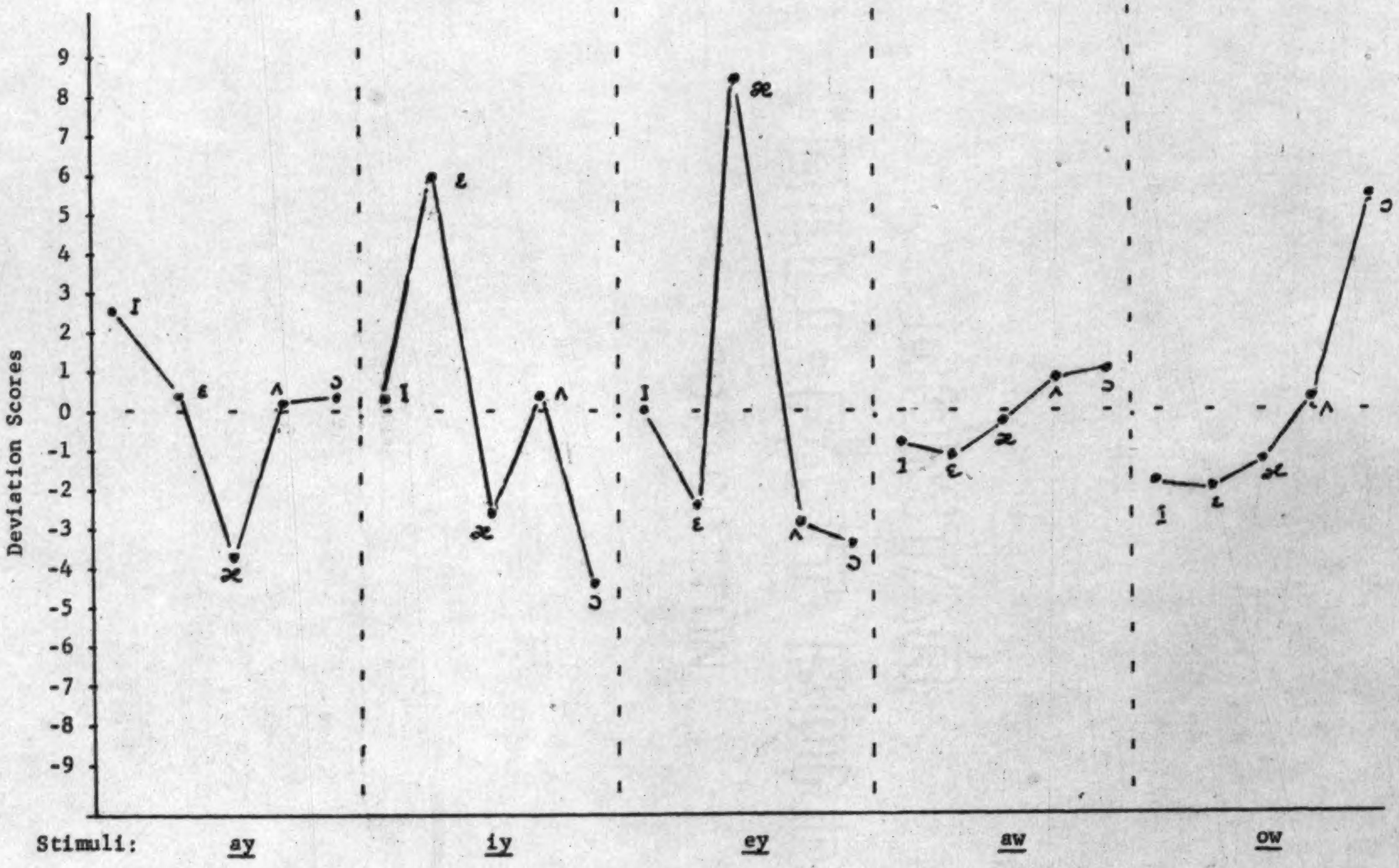

Figure 6A. Stimulus-Actual Response Interaction, Transfer Test, Front Formation

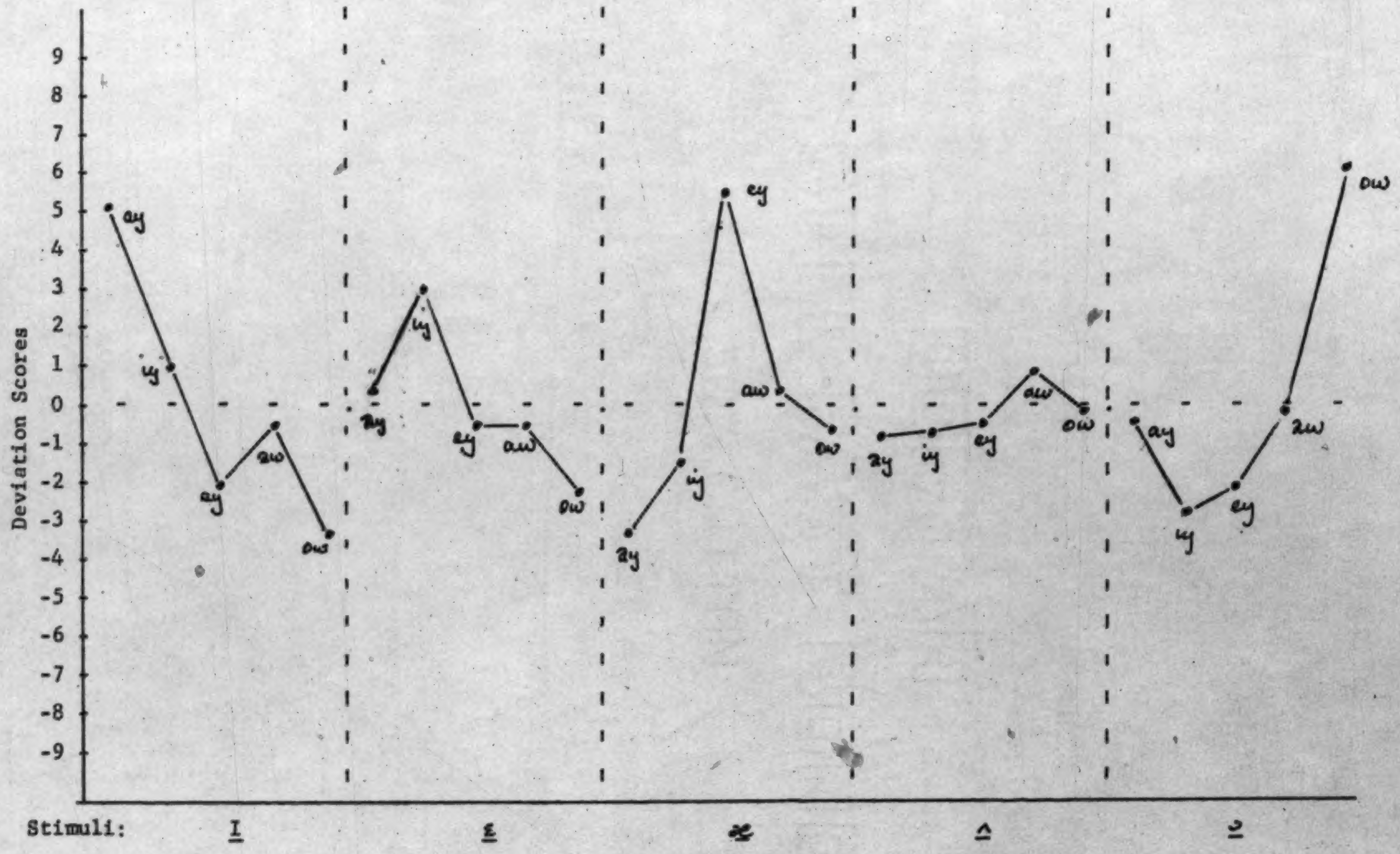

Figure 6B. Stimulus-Actual Response Interaction, Transfer Test, Back Formation